

1. [Preface](#)
2. Welcome to Economics!
 1. [Introduction](#)
 2. [What Is Economics, and Why Is It Important?](#)
 3. [Microeconomics and Macroeconomics](#)
 4. [How Economists Use Theories and Models to Understand Economic Issues](#)
 5. [How Economies Can Be Organized: An Overview of Economic Systems](#)
3. Choice in a World of Scarcity
 1. [Introduction to Choice in a World of Scarcity](#)
 2. [How Individuals Make Choices Based on Their Budget Constraint](#)
 3. [The Production Possibilities Frontier and Social Choices](#)
4. Demand and Supply
 1. [Introduction to Demand and Supply](#)
 2. [Demand, Supply, and Equilibrium in Markets for Goods and Services](#)
 3. [Shifts in Demand and Supply for Goods and Services](#)
 4. [Changes in Equilibrium Price and Quantity: The Four-Step Process](#)
 5. [Price Ceilings and Price Floors](#)
 6. [Demand, Supply, and Efficiency](#)
5. Negative Externalities
 1. [Negative Externalities](#)
 2. [The Economics of Pollution](#)
6. Positive Externalities and Public Goods
 1. [Introduction to Positive Externalities and Public Goods](#)
 2. [How Governments Can Encourage Innovation](#)
 3. [Public Goods](#)
7. The Macroeconomic Perspective

1. [Introduction to the Macroeconomic Perspective](#)
2. [Measuring the Size of the Economy: Gross Domestic Product](#)
3. [Adjusting Nominal Values to Real Values](#)
4. [Tracking Real GDP over Time](#)
5. [Comparing GDP among Countries](#)
6. [How Well GDP Measures the Well-Being of Society](#)
8. Economic Growth
 1. [Introduction to Economic Growth](#)
 2. [The Relatively Recent Arrival of Economic Growth](#)
 3. [Components of Economic Growth](#)
9. Unemployment
 1. [Introduction to Unemployment](#)
 2. [How the Unemployment Rate is Defined and Computed](#)
 3. [Patterns of Unemployment](#)
 4. [What Causes Changes in Unemployment over the Short Run](#)
 5. [What Causes Changes in Unemployment over the Long Run](#)
10. Inflation
 1. [Introduction to Inflation](#)
 2. [Tracking Inflation](#)
 3. [How Changes in the Cost of Living are Measured](#)
 4. [How the U.S. and Other Countries Experience Inflation](#)
 5. [The Confusion Over Inflation](#)
 6. [Indexing and Its Limitations](#)
11. Government Budgets and Fiscal Policy
 1. [Introduction to Government Budgets and Fiscal Policy](#)
 2. [Government Spending](#)
 3. [Taxation](#)
 4. [Federal Deficits and the National Debt](#)

5. [Using Fiscal Policy to Fight Recession, Unemployment, and Inflation](#)
6. [Automatic Stabilizers](#)
7. [Practical Problems with Discretionary Fiscal Policy](#)
8. [The Question of a Balanced Budget](#)
12. The Impacts of Government Borrowing
 1. [Introduction to the Impacts of Government Borrowing](#)
 2. [How Government Borrowing Affects Investment and the Trade Balance](#)
 3. [Fiscal Policy, Investment, and Economic Growth](#)
 4. [How Government Borrowing Affects Private Saving](#)
 5. [Fiscal Policy and the Trade Balance](#)
13. Money and Banking
 1. [Introduction to Money and Banking](#)
 2. [Defining Money by Its Functions](#)
 3. [Measuring Money: Currency, M1, and M2](#)
 4. [The Role of Banks](#)
 5. [How Banks Create Money](#)
14. Monetary Policy and Bank Regulation
 1. [Introduction to Monetary Policy and Bank Regulation](#)
 2. [The Federal Reserve Banking System and Central Banks](#)
 3. [Bank Regulation](#)
 4. [How a Central Bank Executes Monetary Policy](#)
 5. [Monetary Policy and Economic Outcomes](#)
 6. [Pitfalls for Monetary Policy](#)
15. Poverty and Economic Inequality
 1. [Introduction to Poverty and Economic Inequality](#)
 2. [Drawing the Poverty Line](#)
 3. [The Poverty Trap](#)
 4. [The Safety Net](#)
 5. [Income Inequality: Measurement and Causes](#)
 6. [Government Policies to Reduce Income Inequality](#)

16. International Trade

1. [Introduction to International Trade](#)
2. [Absolute and Comparative Advantage](#)
3. [What Happens When a Country Has an Absolute Advantage in All Goods](#)
4. [Intra-industry Trade between Similar Economies](#)
5. [The Benefits of Reducing Barriers to International Trade](#)

17. Globalization and Protectionism

1. [Introduction to Globalization and Protectionism](#)
2. [Protectionism: An Indirect Subsidy from Consumers to Producers](#)
3. [International Trade and Its Effects on Jobs, Wages, and Working Conditions](#)
4. [Arguments in Support of Restricting Imports](#)
5. [How Trade Policy Is Enacted: Globally, Regionally, and Nationally](#)
6. [The Tradeoffs of Trade Policy](#)

18. [The Use of Mathematics in Principles of Economics](#)

Preface

Principles of Economics is designed for a two-semester principles of economics sequence. It is traditional in coverage, including introductory economics content, microeconomics, macroeconomics and international economics. At the same time, the book includes a number of innovative and interactive features designed to enhance student learning. Instructors can also customize the book, adapting it to the approach that works best in their classroom.

Welcome to *Principles of Economics*, an OpenStax resource. This textbook has been created with several goals in mind: accessibility, customization, and student engagement—all while encouraging students toward high levels of academic scholarship. Instructors and students alike will find that this textbook offers a strong foundation in economics in an accessible format.

About OpenStax

OpenStax is a non-profit organization committed to improving student access to quality learning materials. Our free textbooks go through a rigorous editorial publishing process. Our texts are developed and peer-reviewed by educators to ensure they are readable, accurate, and meet the scope and sequence requirements of today's college courses. Unlike traditional textbooks, OpenStax resources live online and are owned by the community of educators using them. Through our partnerships with companies and foundations committed to reducing costs for students, OpenStax is working to improve access to higher education for all. OpenStax is an initiative of Rice University and is made possible through the generous support of several philanthropic foundations.

About OpenStax's Resources

OpenStax resources provide quality academic instruction. Three key features set our materials apart from others: they can be customized by instructors for each class, they are a "living" resource that grows online through contributions from science educators, and they are available free or for minimal cost.

Customization

OpenStax learning resources are designed to be customized for each course. Our textbooks provide a solid foundation on which instructors can build, and our resources are conceived and written with flexibility in mind. Instructors can select the sections most relevant to their curricula and create a textbook that speaks directly to the needs of their classes and student body. Teachers are encouraged to expand on existing examples by adding unique context via geographically localized applications and topical connections.

Principles of Economics can be easily customized using our online platform (<http://cnx.org/content/col11613/>). Simply select the content most relevant to your current semester and create a textbook that speaks directly to the needs of your class. *Principles of Economics* is organized as a collection of sections that can be rearranged, modified, and enhanced through localized examples or to incorporate a specific theme of your course. This customization feature will ensure that your textbook truly reflects the goals of your course. *Principles of Economics* is also available in two volumes, one covering microeconomics and one covering macroeconomics principles.

Curation

To broaden access and encourage community curation, *Principles of Economics* is “open source” licensed under a Creative Commons Attribution (CC-BY) license. The economics community is invited to submit examples, emerging research, and other feedback to enhance and strengthen the material and keep it current and relevant for today’s students. You can submit your suggestions to info@openstaxcollege.org.

Cost

Our textbooks are available for free online, and in low-cost print and e-book editions.

About *Principles of Economics*

Principles of Economics is designed for a two-semester principles of economics sequence. The text has been developed to meet the scope and sequence of most introductory courses. At the same time, the book includes a number of innovative features designed to enhance student learning. Instructors can also customize the book, adapting it to the approach that works best in their classroom.

Coverage and Scope

To develop *Principles of Economics*, we acquired the rights to Timothy Taylor's second edition of *Principles of Economics* and solicited ideas from economics instructors at all levels of higher education, from community colleges to Ph.D.-granting universities. They told us about their courses, students, challenges, resources, and how a textbook can best meet the needs of both instructors and students.

The result is a book that covers the breadth of economics topics and also provides the necessary depth to ensure the course is manageable for instructors and students alike. And to make it more applied, we have incorporated many current topics. We hope students will be interested to know just how far-reaching the recent recession was (and still is), for example, and why there is so much controversy even among economists over the Affordable Care Act (Obamacare). The Keystone Pipeline, Occupy Wall Street, minimum wage debates, and the appointment of the United States' first female Federal Reserve chair, Janet Yellen, are just a few of the other important topics covered.

The pedagogical choices, chapter arrangements, and learning objective fulfillment were developed and vetted with feedback from educators dedicated to the project. They thoroughly read the material and offered critical and detailed commentary. The outcome is a balanced approach to

micro and macro economics, to both Keynesian and classical views, and to the theory and application of economics concepts. New 2015 data are incorporated for topics that range from average U.S. household consumption in Chapter 2 to the total value of all home equity in Chapter 17. Current events are treated in a politically-balanced way as well.

The book is organized into eight main parts:

What is Economics? The first two chapters introduce students to the study of economics with a focus on making choices in a world of scarce resources.

Supply and Demand, Chapters 3 and 4, introduces and explains the first analytical model in economics: supply, demand, and equilibrium, before showing applications in the markets for labor and finance.

The Fundamentals of Microeconomic Theory, Chapters 5 through 10, begins the microeconomics portion of the text, presenting the theories of consumer behavior, production and costs, and the different models of market structure, including some simple game theory.

Microeconomic Policy Issues, Chapters 11 through 18, cover the range of topics in applied micro, framed around the concepts of public goods and positive and negative externalities. Students explore competition and antitrust policies, environmental problems, poverty, income inequality, and other labor market issues. The text also covers information, risk and financial markets, as well as public economy.

The Macroeconomic Perspective and Goals, Chapters 19 through 23, introduces a number of key concepts in macro: economic growth, unemployment and inflation, and international trade and capital flows.

A Framework for Macroeconomic Analysis, Chapters 24 through 26, introduces the principal analytic model in macro, namely the Aggregate Demand/Aggregate Supply Model. The model is then applied to the Keynesian and Neoclassical perspectives. The Expenditure-Output model is fully explained in a stand-alone appendix.

Monetary and Fiscal Policy, Chapters 27 through 31, explains the role of money and the banking system, as well as monetary policy and financial regulation. Then the discussion switches to government deficits and fiscal policy.

International Economics, Chapters 32 through 34, the final part of the text, introduces the international dimensions of economics, including international trade and protectionism.

Chapter 1 Welcome to Economics!
Chapter 2 Choice in a World of Scarcity
Chapter 3 Demand and Supply
Chapter 4 Labor and Financial Markets
Chapter 5 Elasticity
Chapter 6 Consumer Choices
Chapter 7 Cost and Industry Structure
Chapter 8 Perfect Competition
Chapter 9 Monopoly
Chapter 10 Monopolistic Competition and Oligopoly
Chapter 11 Monopoly and Antitrust Policy
Chapter 12 Environmental Protection and Negative Externalities
Chapter 13 Positive Externalities and Public Goods
Chapter 14 Poverty and Economic Inequality
Chapter 15 Issues in Labor Markets: Unions, Discrimination, Immigration
Chapter 16 Information, Risk, and Insurance
Chapter 17 Financial Markets
Chapter 18 Public Economy
Chapter 19 The Macroeconomic Perspective
Chapter 20 Economic Growth
Chapter 21 Unemployment
Chapter 22 Inflation
Chapter 23 The International Trade and Capital Flows
Chapter 24 The Aggregate Demand/Aggregate Supply Model
Chapter 25 The Keynesian Perspective
Chapter 26 The Neoclassical Perspective
Chapter 27 Money and Banking
Chapter 28 Monetary Policy and Bank Regulation
Chapter 29 Exchange Rates and International Capital Flows
Chapter 30 Government Budgets and Fiscal Policy
Chapter 31 The Impacts of Government Borrowing
Chapter 32 Macroeconomic Policy Around the World
Chapter 33 International Trade

Chapter 34 Globalization and Protectionism

Appendix A The Use of Mathematics in Principles of Economics

Appendix B Indifference Curves

Appendix C Present Discounted Value

Appendix D The Expenditure-Output Model

Alternate Sequencing

Principles of Economics was conceived and written to fit a particular topical sequence, but it can be used flexibly to accommodate other course structures. One such potential structure, which will fit reasonably well with the textbook content, is provided. Please consider, however, that the chapters were not written to be completely independent, and that the proposed alternate sequence should be carefully considered for student preparation and textual consistency.

Chapter 1 Welcome to Economics!

Chapter 2 Choice in a World of Scarcity

Chapter 3 Demand and Supply

Chapter 4 Labor and Financial Markets

Chapter 5 Elasticity

Chapter 6 Consumer Choices

Chapter 33 International Trade

Chapter 7 Cost and Industry Structure

Chapter 12 Environmental Protection and Negative Externalities

Chapter 13 Positive Externalities and Public Goods

Chapter 8 Perfect Competition

Chapter 9 Monopoly

Chapter 10 Monopolistic Competition and Oligopoly

Chapter 11 Monopoly and Antitrust Policy

Chapter 14 Poverty and Economic Inequality

Chapter 15 Issues in Labor Markets: Unions, Discrimination, Immigration

Chapter 16 Information, Risk, and Insurance

Chapter 17 Financial Markets

Chapter 18 Public Economy

Chapter 19 The Macroeconomic Perspective

Chapter 20 Economic Growth

Chapter 21 Unemployment
Chapter 22 Inflation
Chapter 23 The International Trade and Capital Flows
Chapter 24 The Aggregate Demand/Aggregate Supply Model
Chapter 25 The Keynesian Perspective
Chapter 26 The Neoclassical Perspective
Chapter 27 Money and Banking
Chapter 28 Monetary Policy and Bank Regulation
Chapter 29 Exchange Rates and International Capital Flows
Chapter 30 Government Budgets and Fiscal Policy
Chapter 31 The Impacts of Government Borrowing
Chapter 32 Macroeconomic Policy Around the World
Chapter 34 Globalization and Protectionism

Appendix A The Use of Mathematics in Principles of Economics
Appendix B Indifference Curves
Appendix C Present Discounted Value
Appendix D The Expenditure-Output Model

Pedagogical Foundation

Throughout the OpenStax version of *Principles of Economics*, you will find new features that engage the students in economic inquiry by taking selected topics a step further. Our features include:

Bring It Home: This added feature is a brief case study, specific to each chapter, which connects the chapter's main topic to the real world. It is broken up into two parts: the first at the beginning of the chapter (in the Intro module) and the second at chapter's end, when students have learned what's necessary to understand the case and "bring home" the chapter's core concepts.

Work It Out: This added feature asks students to work through a generally analytical or computational problem, and guides them step-by-step to find out how its solution is derived.

Clear It Up: This boxed feature, which includes pre-existing features from Taylor's text, addresses common student misconceptions about

the content. Clear It Ups are usually deeper explanations of something in the main body of the text. Each CIU starts with a question. The rest of the feature explains the answer.

Link It Up: This added feature is a very brief introduction to a website that is pertinent to students' understanding and enjoyment of the topic at hand.

Questions for Each Level of Learning

The OpenStax version of *Principles of Economics* further expands on Taylor's original end of chapter materials by offering four types of end-of-module questions for students.

Self-Checks: Are analytical self-assessment questions that appear at the end of each module. They “click-to-reveal” an answer in the web view so students can check their understanding before moving on to the next module. Self-Check questions are not simple look-up questions. They push the student to think a bit beyond what is said in the text. Self-Check questions are designed for formative (rather than summative) assessment. The questions and answers are explained so that students feel like they are being walked through the problem.

Review Questions: Have been retained from Taylor's version, and are simple recall questions from the chapter and are in open-response format (not multiple choice or true/false). The answers can be looked up in the text.

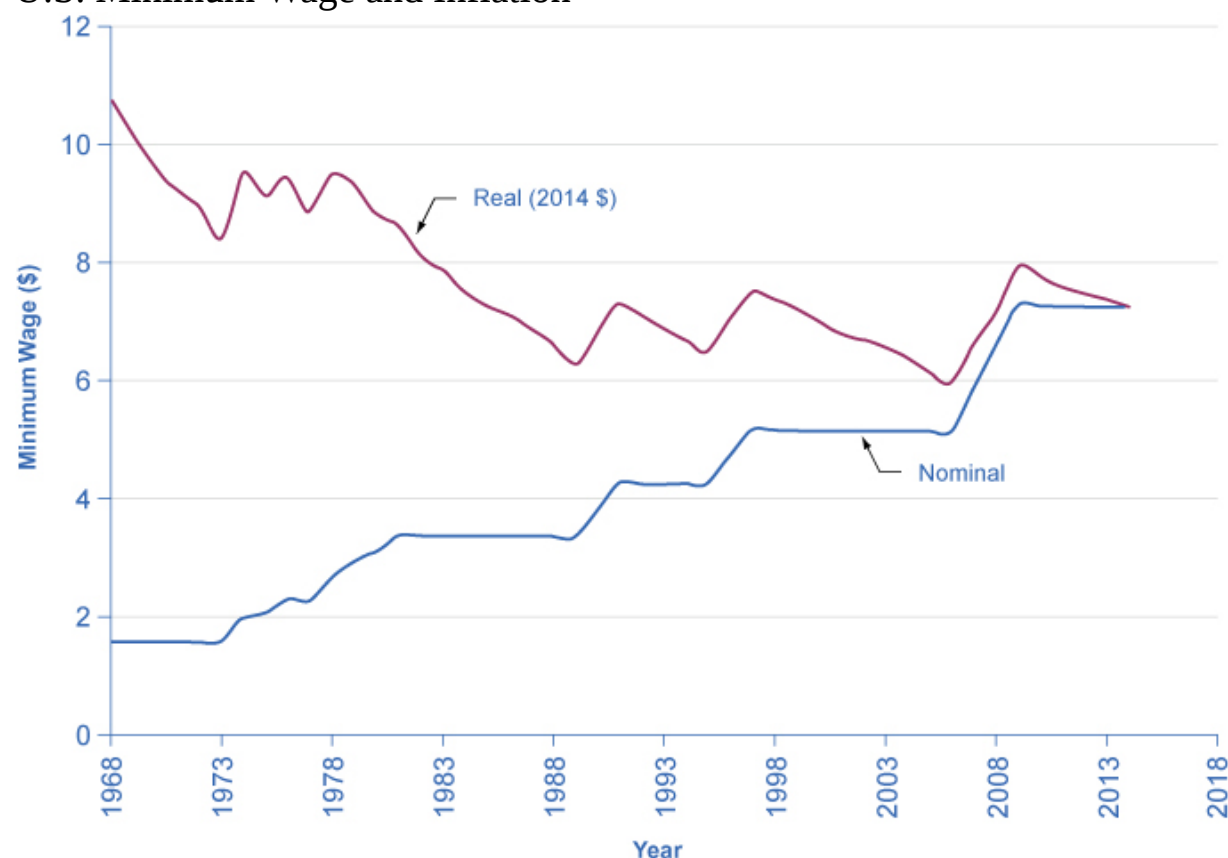
Critical Thinking Questions: Are new higher-level, conceptual questions that ask students to *demonstrate their understanding by applying* what they have learned in different contexts. They ask for outside-the-box thinking, for *reasoning* about the concepts. They push the student to places they wouldn't have thought of going themselves.

Problems: Are exercises that give students additional practice working with the analytic and computational concepts in the module.

Updated Art

Principles of Economics includes an updated art program to better inform today's student, providing the latest data on covered topics.

U.S. Minimum Wage and Inflation



After adjusting for inflation, the federal minimum wage dropped more than 30 percent from 1967 to 2010, even though the nominal figure climbed from \$1.40 to \$7.25 per hour. Increases in the minimum wage in 2007, 2008, and 2009 kept the decline from being worse—as it would have been if the wage had remained the same as it did from 1997 through 2006. (Sources: <http://www.dol.gov/whd/minwage/chart.htm>; <http://data.bls.gov/cgi-bin/surveymost?cu>)

About Our Team

Senior Contributing Author

Timothy Taylor, Macalester College

Timothy Taylor has been writing and teaching about economics for 30 years, and is the Managing Editor of the *Journal of Economic Perspectives*, a post he's held since 1986. He has been a lecturer for The Teaching Company, the University of Minnesota, and the Hubert H. Humphrey Institute of Public Affairs, where students voted him Teacher of the Year in 1997. His writings include numerous pieces for journals such as the *Milken Institute Review* and *The Public Interest*, and he has been an editor on many projects, most notably for the Brookings Institution and the World Bank, where he was Chief Outside Editor for the *World Development Report 1999/2000, Entering the 21st Century: The Changing Development Landscape*. He also blogs four to five times per week at <http://conversableeconomist.blogspot.com>. Timothy Taylor lives near Minneapolis with his wife Kimberley and their three children.

Steven A. Greenlaw, University of Mary Washington

Steven Greenlaw has been teaching principles of economics for more than 30 years. In 1999, he received the Grellet C. Simpson Award for Excellence in Undergraduate Teaching at the University of Mary Washington. He is the author of *Doing Economics: A Guide to Doing and Understanding Economic Research*, as well as a variety of articles on economics pedagogy and instructional technology, published in the *Journal of Economic Education*, the *International Review of Economic Education*, and other outlets. He wrote the module on Quantitative Writing for *Starting Point: Teaching and Learning Economics*, the web portal on best practices in teaching economics. Steven Greenlaw lives in Alexandria, Virginia with his wife Kathy and their three children.

Contributing Authors

Eric Dodge, Hanover College

Cynthia Gamez, University of Texas at El Paso

Andres Jauregui, Columbus State University

Diane Keenan, Cerritos College

Dan MacDonald, California State University San Bernardino

Amyaz Moledina, The College of Wooster

Craig Richardson, Winston-Salem State University
David Shapiro, Pennsylvania State University
Ralph Sonenshine, American University

Expert Reviewers

Bryan Aguiar, Northwest Arkansas Community College
Basil Al Hashimi, Mesa Community College
Emil Berendt, Mount St. Mary's University
Zena Buser, Adams State University
Douglas Campbell, The University of Memphis
Sanjukta Chaudhuri, University of Wisconsin - Eau Claire
Xueyu Cheng, Alabama State University
Robert Cunningham, Alma College
Rosa Lea Danielson, College of DuPage
Steven Deloach, Elon University
Debbie Evercloud, University of Colorado Denver
Sal Figueras, Hudson County Community College
Reza Ghorashi, Richard Stockton College of New Jersey
Robert Gillette, University of Kentucky
Shaomin Huang, Lewis-Clark State College
George Jones, University of Wisconsin-Rock County
Charles Kroncke, College of Mount St. Joseph
Teresa Laughlin, Palomar Community College
Carlos Liard-Muriente, Central Connecticut State University
Heather Luea, Kansas State University
Steven Lugauer, University of Notre Dame
William Mosher, Nashua Community College
Michael Netta, Hudson County Community College
Nick Noble, Miami University
Joe Nowakowski, Muskingum University
Shawn Osell, University of Wisconsin, Superior
Mark Owens, Middle Tennessee State University
Sonia Pereira, Barnard College
Brian Peterson, Central College
Jennifer Platania, Elon University

Robert Rycroft, University of Mary Washington
Adrienne Sachse, Florida State College at Jacksonville
Hans Schumann, Texas AM
Gina Shamshak, Goucher College
Chris Warburton, John Jay College of Criminal Justice, CUNY
Mark Witte, Northwestern
Chiou-nan Yeh, Alabama State University

Ancillaries

OpenStax projects offer an array of ancillaries for students and instructors. Please visit <http://openstaxcollege.org> and view the learning resources for this title.

Introduction
class="introduction"
Do You Use Facebook?

Economics is
greatly
impacted by
how well
information
travels through
society. Today,
social media
giants Twitter,
Facebook, and
Instagram are
major forces
on the
information
super highway.
(Credit: Johan
Larsson/Flickr
)

**Note:****Decisions ... Decisions in the Social Media Age**

To post or not to post? Every day we are faced with a myriad of decisions, from what to have for breakfast, to which route to take to class, to the more complex—“Should I double major and add possibly another semester of study to my education?” Our response to these choices depends on the information we have available at any given moment; information economists call “imperfect” because we rarely have all the data we need to make perfect decisions. Despite the lack of perfect information, we still make hundreds of decisions a day.

And now, we have another avenue in which to gather information—social media. Outlets like Facebook and Twitter are altering the process by which we make choices, how we spend our time, which movies we see, which products we buy, and more. How many of you chose a university without checking out its Facebook page or Twitter stream first for information and feedback?

As you will see in this course, what happens in economics is affected by how well and how fast information is disseminated through a society, such as how quickly information travels through Facebook. “Economists love nothing better than when deep and liquid markets operate under conditions of perfect information,” says Jessica Irvine, National Economics Editor for News Corp Australia.

This leads us to the topic of this chapter, an introduction to the world of making decisions, processing information, and understanding behavior in markets—the world of economics. Each chapter in this book will start with a discussion about current (or sometimes past) events and revisit it at chapter’s end—to “bring home” the concepts in play.

Note:**Introduction**

In this chapter, you will learn about:

- What Is Economics, and Why Is It Important?
- Microeconomics and Macroeconomics
- How Economists Use Theories and Models to Understand Economic Issues
- How Economies Can Be Organized: An Overview of Economic Systems

What is economics and why should you spend your time learning it? After all, there are other disciplines you could be studying, and other ways you could be spending your time. As the Bring it Home feature just mentioned, making choices is at the heart of what economists study, and your decision to take this course is as much an economic decision as anything else.

Economics is probably not what you think. It is not primarily about money or finance. It is not primarily about business. It is not mathematics. What is it then? It is both a subject area and a way of viewing the world.

What Is Economics, and Why Is It Important?

By the end of this section, you will be able to:

- Discuss the importance of studying economics
- Explain the relationship between production and division of labor
- Evaluate the significance of scarcity

Economics is the study of how humans make decisions in the face of scarcity. These can be individual decisions, family decisions, business decisions or societal decisions. If you look around carefully, you will see that scarcity is a fact of life. **Scarcity** means that human wants for goods, services and resources exceed what is available. Resources, such as labor, tools, land, and raw materials are necessary to produce the goods and services we want but they exist in limited supply. Of course, the ultimate scarce resource is time- everyone, rich or poor, has just 24 hours in the day to try to acquire the goods they want. At any point in time, there is only a finite amount of resources available.

Think about it this way: In 2015 the labor force in the United States contained over 158.6 million workers, according to the U.S. Bureau of Labor Statistics. Similarly, the total area of the United States is 3,794,101 square miles. These are large numbers for such crucial resources, however, they are limited. Because these resources are limited, so are the numbers of goods and services we produce with them. Combine this with the fact that human wants seem to be virtually infinite, and you can see why scarcity is a problem.

Scarcity of Resources



Homeless people are a stark reminder that scarcity of resources is real. (Credit: “daveynin”/Flickr Creative Commons)

If you still do not believe that scarcity is a problem, consider the following: Does everyone need food to eat? Does everyone need a decent place to live? Does everyone have access to healthcare? In every country in the world, there are people who are hungry, homeless (for example, those who call park benches their beds, as shown in [\[link\]](#)), and in need of healthcare, just to focus on a few critical goods and services. Why is this the case? It is because of scarcity. Let’s delve into the concept of scarcity a little deeper, because it is crucial to understanding economics.

The Problem of Scarcity

Think about all the things you consume: food, shelter, clothing, transportation, healthcare, and entertainment. How do you acquire those items? You do not produce them yourself. You buy them. How do you afford the things you buy? You work for pay. Or if you do not, someone else does on your behalf. Yet most of us never have enough to buy all the things we want. This is because of scarcity. So how do we solve it?

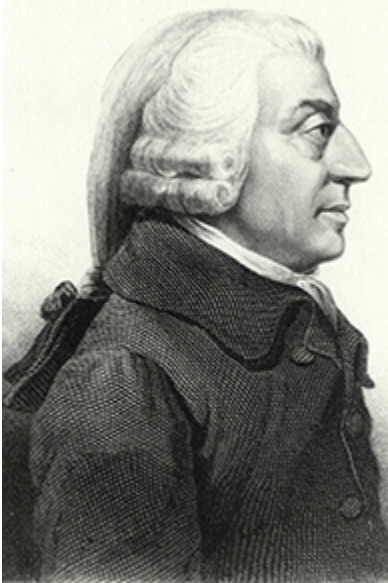
Note:

Visit this [website](#) to read about how the United States is dealing with scarcity in resources.



Every society, at every level, must make choices about how to use its resources. Families must decide whether to spend their money on a new car or a fancy vacation. Towns must choose whether to put more of the budget into police and fire protection or into the school system. Nations must decide whether to devote more funds to national defense or to protecting the environment. In most cases, there just isn't enough money in the budget to do everything. So why do we not each just produce all of the things we consume? The simple answer is most of us do not know how, but that is not the main reason. (When you study economics, you will discover that the obvious choice is not always the right answer—or at least the complete answer. Studying economics teaches you to think in a different of way.) Think back to pioneer days, when individuals knew how to do so much more than we do today, from building their homes, to growing their crops, to hunting for food, to repairing their equipment. Most of us do not know how to do all—or any—of those things. It is not because we could not learn. Rather, we do not have to. The reason why is something called *the division and specialization of labor*, a production innovation first put forth by Adam Smith, [\[link\]](#), in his book, *The Wealth of Nations*.

Adam Smith



Adam Smith
introduced the idea
of dividing labor
into discrete tasks.

(Credit:
Wikimedia
Commons)

The Division of and Specialization of Labor

The formal study of economics began when Adam Smith (1723–1790) published his famous book *The Wealth of Nations* in 1776. Many authors had written on economics in the centuries before Smith, but he was the first to address the subject in a comprehensive way. In the first chapter, Smith introduces the **division of labor**, which means that the way a good or service is produced is divided into a number of tasks that are performed by different workers, instead of all the tasks being done by the same person.

To illustrate the division of labor, Smith counted how many tasks went into making a pin: drawing out a piece of wire, cutting it to the right length, straightening it, putting a head on one end and a point on the other, and

packaging pins for sale, to name just a few. Smith counted 18 distinct tasks that were often done by different people—all for a pin, believe it or not!

Modern businesses divide tasks as well. Even a relatively simple business like a restaurant divides up the task of serving meals into a range of jobs like top chef, sous chefs, less-skilled kitchen help, servers to wait on the tables, a greeter at the door, janitors to clean up, and a business manager to handle paychecks and bills—not to mention the economic connections a restaurant has with suppliers of food, furniture, kitchen equipment, and the building where it is located. A complex business like a large manufacturing factory, such as the shoe factory shown in [\[link\]](#), or a hospital can have hundreds of job classifications.

Division of Labor



Workers on an assembly line are an example of the divisions of labor. (Credit: Nina Hale/Flickr Creative Commons)

Why the Division of Labor Increases Production

When the tasks involved with producing a good or service are divided and subdivided, workers and businesses can produce a greater quantity of output. In his observations of pin factories, Smith observed that one worker

alone might make 20 pins in a day, but that a small business of 10 workers (some of whom would need to do two or three of the 18 tasks involved with pin-making), could make 48,000 pins in a day. How can a group of workers, each specializing in certain tasks, produce so much more than the same number of workers who try to produce the entire good or service by themselves? Smith offered three reasons.

First, **specialization** in a particular small job allows workers to focus on the parts of the production process where they have an advantage. (In later chapters, we will develop this idea by discussing comparative advantage.) People have different skills, talents, and interests, so they will be better at some jobs than at others. The particular advantages may be based on educational choices, which are in turn shaped by interests and talents. Only those with medical degrees qualify to become doctors, for instance. For some goods, specialization will be affected by geography—it is easier to be a wheat farmer in North Dakota than in Florida, but easier to run a tourist hotel in Florida than in North Dakota. If you live in or near a big city, it is easier to attract enough customers to operate a successful dry cleaning business or movie theater than if you live in a sparsely populated rural area. Whatever the reason, if people specialize in the production of what they do best, they will be more productive than if they produce a combination of things, some of which they are good at and some of which they are not.

Second, workers who specialize in certain tasks often learn to produce more quickly and with higher quality. This pattern holds true for many workers, including assembly line laborers who build cars, stylists who cut hair, and doctors who perform heart surgery. In fact, specialized workers often know their jobs well enough to suggest innovative ways to do their work faster and better.

A similar pattern often operates within businesses. In many cases, a business that focuses on one or a few products (sometimes called its “core competency”) is more successful than firms that try to make a wide range of products.

Third, specialization allows businesses to take advantage of **economies of scale**, which means that for many goods, as the level of production increases, the average cost of producing each individual unit declines. For

example, if a factory produces only 100 cars per year, each car will be quite expensive to make on average. However, if a factory produces 50,000 cars each year, then it can set up an assembly line with huge machines and workers performing specialized tasks, and the average cost of production per car will be lower. The ultimate result of workers who can focus on their preferences and talents, learn to do their specialized jobs better, and work in larger organizations is that society as a whole can produce and consume far more than if each person tried to produce all of their own goods and services. The division and specialization of labor has been a force against the problem of scarcity.

Trade and Markets

Specialization only makes sense, though, if workers can use the pay they receive for doing their jobs to purchase the other goods and services that they need. In short, specialization requires trade.

You do not have to know anything about electronics or sound systems to play music—you just buy an iPod or MP3 player, download the music and listen. You do not have to know anything about artificial fibers or the construction of sewing machines if you need a jacket—you just buy the jacket and wear it. You do not need to know anything about internal combustion engines to operate a car—you just get in and drive. Instead of trying to acquire all the knowledge and skills involved in producing all of the goods and services that you wish to consume, the market allows you to learn a specialized set of skills and then use the pay you receive to buy the goods and services you need or want. This is how our modern society has evolved into a strong economy.

Why Study Economics?

Now that we have gotten an overview on what economics studies, let's quickly discuss why you are right to study it. Economics is not primarily a collection of facts to be memorized, though there are plenty of important concepts to be learned. Instead, economics is better thought of as a collection of questions to be answered or puzzles to be worked out. Most important, economics provides the tools to work out those puzzles. If you

have yet to be bitten by the economics “bug,” there are other reasons why you should study economics.

- Virtually every major problem facing the world today, from global warming, to world poverty, to the conflicts in Syria, Afghanistan, and Somalia, has an economic dimension. If you are going to be part of solving those problems, you need to be able to understand them. Economics is crucial.
- It is hard to overstate the importance of economics to good citizenship. You need to be able to vote intelligently on budgets, regulations, and laws in general. When the U.S. government came close to a standstill at the end of 2012 due to the “fiscal cliff,” what were the issues involved? Did you know?
- A basic understanding of economics makes you a well-rounded thinker. When you read articles about economic issues, you will understand and be able to evaluate the writer’s argument. When you hear classmates, co-workers, or political candidates talking about economics, you will be able to distinguish between common sense and nonsense. You will find new ways of thinking about current events and about personal and business decisions, as well as current events and politics.

The study of economics does not dictate the answers, but it can illuminate the different choices.

Key Concepts and Summary

Economics seeks to solve the problem of scarcity, which is when human wants for goods and services exceed the available supply. A modern economy displays a division of labor, in which people earn income by specializing in what they produce and then use that income to purchase the products they need or want. The division of labor allows individuals and firms to specialize and to produce more for several reasons: a) It allows the agents to focus on areas of advantage due to natural factors and skill levels; b) It encourages the agents to learn and invent; c) It allows agents to take advantage of economies of scale. Division and specialization of labor only work when individuals can purchase what they do not produce in markets.

Learning about economics helps you understand the major problems facing the world today, prepares you to be a good citizen, and helps you become a well-rounded thinker.

Self-Check Questions

Exercise:

Problem: What is scarcity? Can you think of two causes of scarcity?

Solution:

Scarcity means human wants for goods and services exceed the available supply. Supply is limited because resources are limited. Demand, however, is virtually unlimited. Whatever the supply, it seems human nature to want more.

Exercise:

Problem:

Residents of the town of Smithfield like to consume hams, but each ham requires 10 people to produce it and takes a month. If the town has a total of 100 people, what is the maximum amount of ham the residents can consume in a month?

Solution:

$100 \text{ people} / 10 \text{ people per ham} = \text{a maximum of 10 hams per month}$ if all residents produce ham. Since consumption is limited by production, the maximum number of hams residents could consume per month is 10.

Exercise:

Problem:

A consultant works for \$200 per hour. She likes to eat vegetables, but is not very good at growing them. Why does it make more economic sense for her to spend her time at the consulting job and shop for her vegetables?

Solution:

She is very productive at her consulting job, but not very productive growing vegetables. Time spent consulting would produce far more income than it what she could save growing her vegetables using the same amount of time. So on purely economic grounds, it makes more sense for her to maximize her income by applying her labor to what she does best (i.e. specialization of labor).

Exercise:**Problem:**

A computer systems engineer could paint his house, but it makes more sense for him to hire a painter to do it. Explain why.

Solution:

The engineer is better at computer science than at painting. Thus, his time is better spent working for pay at his job and paying a painter to paint his house. Of course, this assumes he does not paint his house for fun!

Review Questions**Exercise:****Problem:**

Give the three reasons that explain why the division of labor increases an economy's level of production.

Exercise:

Problem: What are three reasons to study economics?

Critical Thinking Questions

Exercise:

Problem:

Suppose you have a team of two workers: one is a baker and one is a chef. Explain why the kitchen can produce more meals in a given period of time if each worker specializes in what they do best than if each worker tries to do everything from appetizer to dessert.

Exercise:

Problem: Why would division of labor without trade not work?

Exercise:

Problem:

Can you think of any examples of *free* goods, that is, goods or services that are not scarce?

References

Bureau of Labor Statistics, U.S. Department of Labor. 2015. "The Employment Situation—February 2015." Accessed March 27, 2015.
<http://www.bls.gov/news.release/pdf/empstat.pdf>.

Williamson, Lisa. "US Labor Market in 2012." *Bureau of Labor Statistics*. Accessed December 1, 2013.
<http://www.bls.gov/opub/mlr/2013/03/art1full.pdf>.

Glossary

division of labor

the way in which the work required to produce a good or service is divided into tasks performed by different workers

economics

the study of how humans make choices under conditions of scarcity

economies of scale

when the average cost of producing each individual unit declines as total output increases

scarcity

when human wants for goods and services exceed the available supply

specialization

when workers or firms focus on particular tasks for which they are well-suited within the overall production process

Microeconomics and Macroeconomics

By the end of this section, you will be able to:

- Describe microeconomics
- Describe macroeconomics
- Contrast monetary policy and fiscal policy

Economics is concerned with the well-being of *all* people, including those with jobs and those without jobs, as well as those with high incomes and those with low incomes. Economics acknowledges that production of useful goods and services can create problems of environmental pollution. It explores the question of how investing in education helps to develop workers' skills. It probes questions like how to tell when big businesses or big labor unions are operating in a way that benefits society as a whole and when they are operating in a way that benefits their owners or members at the expense of others. It looks at how government spending, taxes, and regulations affect decisions about production and consumption.

It should be clear by now that economics covers a lot of ground. That ground can be divided into two parts: **Microeconomics** focuses on the actions of individual agents within the economy, like households, workers, and businesses; **Macroeconomics** looks at the economy as a whole. It focuses on broad issues such as growth of production, the number of unemployed people, the inflationary increase in prices, government deficits, and levels of exports and imports. Microeconomics and macroeconomics are not separate subjects, but rather complementary perspectives on the overall subject of the economy.

To understand why both microeconomic and macroeconomic perspectives are useful, consider the problem of studying a biological ecosystem like a lake. One person who sets out to study the lake might focus on specific topics: certain kinds of algae or plant life; the characteristics of particular fish or snails; or the trees surrounding the lake. Another person might take an overall view and instead consider the entire ecosystem of the lake from top to bottom; what eats what, how the system stays in a rough balance, and what environmental stresses affect this balance. Both approaches are useful, and both examine the same lake, but the viewpoints are different. In a

similar way, both microeconomics and macroeconomics study the same economy, but each has a different viewpoint.

Whether you are looking at lakes or economics, the micro and the macro insights should blend with each other. In studying a lake, the micro insights about particular plants and animals help to understand the overall food chain, while the macro insights about the overall food chain help to explain the environment in which individual plants and animals live.

In economics, the micro decisions of individual businesses are influenced by whether the macroeconomy is healthy; for example, firms will be more likely to hire workers if the overall economy is growing. In turn, the performance of the macroeconomy ultimately depends on the microeconomic decisions made by individual households and businesses.

Microeconomics

What determines how households and individuals spend their budgets? What combination of goods and services will best fit their needs and wants, given the budget they have to spend? How do people decide whether to work, and if so, whether to work full time or part time? How do people decide how much to save for the future, or whether they should borrow to spend beyond their current means?

What determines the products, and how many of each, a firm will produce and sell? What determines what prices a firm will charge? What determines how a firm will produce its products? What determines how many workers it will hire? How will a firm finance its business? When will a firm decide to expand, downsize, or even close? In the microeconomic part of this book, we will learn about the theory of consumer behavior and the theory of the firm.

Macroeconomics

What determines the level of economic activity in a society? In other words, what determines how many goods and services a nation actually produces? What determines how many jobs are available in an economy? What

determines a nation's standard of living? What causes the economy to speed up or slow down? What causes firms to hire more workers or to lay workers off? Finally, what causes the economy to grow over the long term?

An economy's macroeconomic health can be defined by a number of goals: growth in the standard of living, low unemployment, and low inflation, to name the most important. How can macroeconomic policy be used to pursue these goals? **Monetary policy**, which involves policies that affect bank lending, interest rates, and financial capital markets, is conducted by a nation's central bank. For the United States, this is the Federal Reserve. **Fiscal policy**, which involves government spending and taxes, is determined by a nation's legislative body. For the United States, this is the Congress and the executive branch, which originates the federal budget. These are the main tools the government has to work with. Americans tend to expect that government can fix whatever economic problems we encounter, but to what extent is that expectation realistic? These are just some of the issues that will be explored in the macroeconomic chapters of this book.

Key Concepts and Summary

Microeconomics and macroeconomics are two different perspectives on the economy. The microeconomic perspective focuses on parts of the economy: individuals, firms, and industries. The macroeconomic perspective looks at the economy as a whole, focusing on goals like growth in the standard of living, unemployment, and inflation. Macroeconomics has two types of policies for pursuing these goals: monetary policy and fiscal policy.

Self-Check Questions

Exercise:

Problem:

What would be another example of a “system” in the real world that could serve as a metaphor for micro and macroeconomics?

Solution:

There are many physical systems that would work, for example, the study of planets (micro) in the solar system (macro), or solar systems (micro) in the galaxy (macro).

Review Questions**Exercise:****Problem:**

What is the difference between microeconomics and macroeconomics?

Exercise:

Problem: What are examples of individual economic agents?

Exercise:

Problem: What are the three main goals of macroeconomics?

Critical Thinking Questions**Exercise:****Problem:**

A balanced federal budget and a balance of trade are considered secondary goals of macroeconomics, while growth in the standard of living (for example) is considered a primary goal. Why do you think that is so?

Exercise:

Problem:

Macroeconomics is an aggregate of what happens at the microeconomic level. Would it be possible for what happens at the macro level to differ from how economic agents would react to some stimulus at the micro level? *Hint:* Think about the behavior of crowds.

Glossary

fiscal policy

economic policies that involve government spending and taxes

macroeconomics

the branch of economics that focuses on broad issues such as growth, unemployment, inflation, and trade balance.

microeconomics

the branch of economics that focuses on actions of particular agents within the economy, like households, workers, and business firms

monetary policy

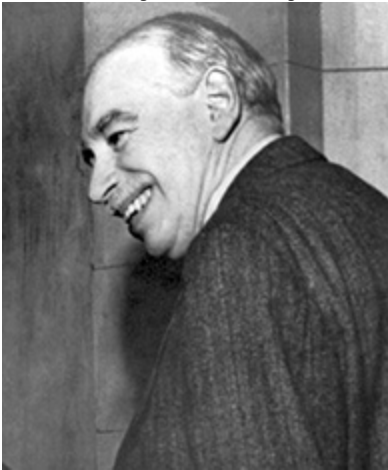
policy that involves altering the level of interest rates, the availability of credit in the economy, and the extent of borrowing

How Economists Use Theories and Models to Understand Economic Issues

By the end of this section, you will be able to:

- Interpret a circular flow diagram
- Explain the importance of economic theories and models
- Describe goods and services markets and labor markets

John Maynard Keynes



One of the most
influential
economists in
modern times was
John Maynard
Keynes. (Credit:
Wikimedia
Commons)

John Maynard Keynes (1883–1946), one of the greatest economists of the twentieth century, pointed out that economics is not just a subject area but also a way of thinking. Keynes, shown in [\[link\]](#), famously wrote in the introduction to a fellow economist's book: “[Economics] is a method rather than a doctrine, an apparatus of the mind, a technique of thinking, which

helps its possessor to draw correct conclusions.” In other words, economics teaches you how to think, not what to think.

Note:

Watch this [video](#) about John Maynard Keynes and his influence on economics.



Economists see the world through a different lens than anthropologists, biologists, classicists, or practitioners of any other discipline. They analyze issues and problems with economic theories that are based on particular assumptions about human behavior, that are different than the assumptions an anthropologist or psychologist might use. A **theory** is a simplified representation of how two or more variables interact with each other. The purpose of a theory is to take a complex, real-world issue and simplify it down to its essentials. If done well, this enables the analyst to understand the issue and any problems around it. A good theory is simple enough to be understood, while complex enough to capture the key features of the object or situation being studied.

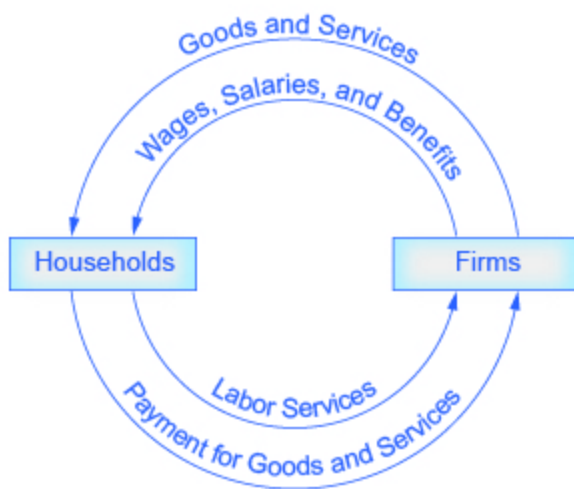
Sometimes economists use the term **model** instead of theory. Strictly speaking, a theory is a more abstract representation, while a model is more applied or empirical representation. Models are used to test theories, but for this course we will use the terms interchangeably.

For example, an architect who is planning a major office building will often build a physical model that sits on a tabletop to show how the entire city

block will look after the new building is constructed. Companies often build models of their new products, which are more rough and unfinished than the final product will be, but can still demonstrate how the new product will work.

A good model to start with in economics is the **circular flow diagram**, which is shown in [\[link\]](#). It pictures the economy as consisting of two groups—households and firms—that interact in two markets: the **goods and services market** in which firms sell and households buy and the **labor market** in which households sell labor to business firms or other employees.

The Circular Flow Diagram



The circular flow diagram shows how households and firms interact in the goods and services market, and in the labor market. The direction of the arrows shows that in the goods and services market, households receive goods and services and pay firms for them. In the labor market, households provide labor and receive payment from firms through wages, salaries, and benefits.

Of course, in the real world, there are many different markets for goods and services and markets for many different types of labor. The circular flow diagram simplifies this to make the picture easier to grasp. In the diagram, firms produce goods and services, which they sell to households in return for revenues. This is shown in the outer circle, and represents the two sides of the product market (for example, the market for goods and services) in which households demand and firms supply. Households sell their labor as workers to firms in return for wages, salaries and benefits. This is shown in the inner circle and represents the two sides of the labor market in which households supply and firms demand.

This version of the circular flow model is stripped down to the essentials, but it has enough features to explain how the product and labor markets work in the economy. We could easily add details to this basic model if we wanted to introduce more real-world elements, like financial markets, governments, and interactions with the rest of the globe (imports and exports).

Economists carry a set of theories in their heads like a carpenter carries around a toolkit. When they see an economic issue or problem, they go through the theories they know to see if they can find one that fits. Then they use the theory to derive insights about the issue or problem. In economics, theories are expressed as diagrams, graphs, or even as mathematical equations. (Do not worry. In this course, we will mostly use graphs.) Economists do not figure out the answer to the problem first and then draw the graph to illustrate. Rather, they use the graph of the theory to help them figure out the answer. Although at the introductory level, you can sometimes figure out the right answer without applying a model, if you keep studying economics, before too long you will run into issues and problems that you will need to graph to solve. Both micro and macroeconomics are explained in terms of theories and models. The most well-known theories are probably those of supply and demand, but you will learn a number of others.

Key Concepts and Summary

Economists analyze problems differently than do other disciplinary experts. The main tools economists use are economic theories or models. A theory is not an illustration of the answer to a problem. Rather, a theory is a tool for determining the answer.

Self-Check Questions

Exercise:

Problem:

Suppose we extend the circular flow model to add imports and exports. Copy the circular flow diagram onto a sheet of paper and then add a foreign country as a third agent. Draw a rough sketch of the flows of imports, exports, and the payments for each on your diagram.

Solution:

Draw a box outside the original circular flow to represent the foreign country. Draw an arrow from the foreign country to firms, to represents imports. Draw an arrow in the reverse direction representing payments for imports. Draw an arrow from firms to the foreign country to represent exports. Draw an arrow in the reverse direction to represent payments for imports.

Exercise:

Problem:

What is an example of a problem in the world today, not mentioned in the chapter, that has an economic dimension?

Solution:

There are many such problems. Consider the AIDS epidemic. Why are so few AIDS patients in Africa and Southeast Asia treated with the same drugs that are effective in the United States and Europe? It is because neither those patients nor the countries in which they live have the resources to purchase the same drugs.

Review Questions

Exercise:

Problem: How did John Maynard Keynes define economics?

Exercise:

Problem:

Are households primarily buyers or sellers in the goods and services market? In the labor market?

Exercise:

Problem:

Are firms primarily buyers or sellers in the goods and services market? In the labor market?

Critical Thinking Questions

Exercise:

Problem:

Why is it unfair or meaningless to criticize a theory as “unrealistic?”

Exercise:

Problem:

Suppose, as an economist, you are asked to analyze an issue unlike anything you have ever done before. Also, suppose you do not have a specific model for analyzing that issue. What should you do? *Hint:* What would a carpenter do in a similar situation?

Glossary

circular flow diagram

a diagram that views the economy as consisting of households and firms interacting in a goods and services market and a labor market

goods and services market

a market in which firms are sellers of what they produce and households are buyers

labor market

the market in which households sell their labor as workers to business firms or other employers

model

see theory

theory

a representation of an object or situation that is simplified while including enough of the key features to help us understand the object or situation

How Economies Can Be Organized: An Overview of Economic Systems

By the end of this section, you will be able to:

- Contrast traditional economies, command economies, and market economies
- Explain gross domestic product (GDP)
- Assess the importance and effects of globalization

Think about what a complex system a modern economy is. It includes all production of goods and services, all buying and selling, all employment. The economic life of every individual is interrelated, at least to a small extent, with the economic lives of thousands or even millions of other individuals. Who organizes and coordinates this system? Who insures that, for example, the number of televisions a society provides is the same as the amount it needs and wants? Who insures that the right number of employees work in the electronics industry? Who insures that televisions are produced in the best way possible? How does it all get done?

There are at least three ways societies have found to organize an economy. The first is the **traditional economy**, which is the oldest economic system and can be found in parts of Asia, Africa, and South America. Traditional economies organize their economic affairs the way they have always done (i.e., tradition). Occupations stay in the family. Most families are farmers who grow the crops they have always grown using traditional methods. What you produce is what you get to consume. Because things are driven by tradition, there is little economic progress or development.

A Command Economy



Ancient Egypt was an example of a command economy. (Credit: Jay Bergesen/Flickr Creative Commons)

Command economies are very different. In a **command economy**, economic effort is devoted to goals passed down from a ruler or ruling class. Ancient Egypt was a good example: a large part of economic life was devoted to building pyramids, like those shown in [\[link\]](#), for the pharaohs. Medieval manor life is another example: the lord provided the land for growing crops and protection in the event of war. In return, vassals provided labor and soldiers to do the lord's bidding. In the last century, communism emphasized command economies.

In a command economy, the government decides what goods and services will be produced and what prices will be charged for them. The government decides what methods of production will be used and how much workers will be paid. Many necessities like healthcare and education are provided for free. Currently, Cuba and North Korea have command economies.

A Market Economy



Nothing says
“market” more
than The New
York Stock
Exchange. (Credit:
Erik Drost/Flickr
Creative
Commons)

Although command economies have a very centralized structure for economic decisions, market economies have a very decentralized structure. A **market** is an institution that brings together buyers and sellers of goods or services, who may be either individuals or businesses. The New York Stock Exchange, shown in [\[link\]](#), is a prime example of market in which buyers and sellers are brought together. In a **market economy**, decision-making is decentralized. Market economies are based on **private enterprise**: the means of production (resources and businesses) are owned and operated by private individuals or groups of private individuals. Businesses supply goods and services based on demand. (In a command economy, by contrast, resources and businesses are owned by the government.) What goods and services are supplied depends on what is demanded. A person’s income is based on his or her ability to convert resources (especially labor) into something that society values. The more society values the person’s output, the higher the income (think Lady Gaga

or LeBron James). In this scenario, economic decisions are determined by market forces, not governments.

Most economies in the real world are mixed; they combine elements of command and market (and even traditional) systems. The U.S. economy is positioned toward the market-oriented end of the spectrum. Many countries in Europe and Latin America, while primarily market-oriented, have a greater degree of government involvement in economic decisions than does the U.S. economy. China and Russia, while they are closer to having a market-oriented system now than several decades ago, remain closer to the command economy end of the spectrum. A rich resource of information about countries and their economies can be found on the Heritage Foundation's website, as the following Clear It Up feature discusses.

Note:

What countries are considered economically free?

Who is in control of economic decisions? Are people free to do what they want and to work where they want? Are businesses free to produce when they want and what they choose, and to hire and fire as they wish? Are banks free to choose who will receive loans? Or does the government control these kinds of choices? Each year, researchers at the Heritage Foundation and the *Wall Street Journal* look at 50 different categories of economic freedom for countries around the world. They give each nation a score based on the extent of economic freedom in each category.

The 2015 Heritage Foundation's Index of Economic Freedom report ranked 178 countries around the world: some examples of the most free and the least free countries are listed in [\[link\]](#). Several countries were not ranked because of extreme instability that made judgments about economic freedom impossible. These countries include Afghanistan, Iraq, Syria, and Somalia.

The assigned rankings are inevitably based on estimates, yet even these rough measures can be useful for discerning trends. In 2015, 101 of the 178 included countries shifted toward greater economic freedom, although 77 of the countries shifted toward less economic freedom. In recent

decades, the overall trend has been a *higher level of economic freedom around the world*.

Most Economic Freedom	Least Economic Freedom
1. Hong Kong	167. Timor-Leste
2. Singapore	168. Democratic Republic of Congo
3. New Zealand	169. Argentina
4. Australia	170. Republic of Congo
5. Switzerland	171. Iran
6. Canada	172. Turkmenistan
7. Chile	173. Equatorial Guinea
8. Estonia	174. Eritrea
9. Ireland	175. Zimbabwe
10. Mauritius	176. Venezuela
11. Denmark	177. Cuba
12. United States	178. North Korea

Economic Freedoms, 2015(Source: The Heritage Foundation, 2015 Index of Economic Freedom, Country Rankings, <http://www.heritage.org/index/ranking>)

Regulations: The Rules of the Game

Markets and government regulations are always entangled. There is no such thing as an absolutely free market. Regulations always define the “rules of the game” in the economy. Economies that are primarily market-oriented have fewer regulations—ideally just enough to maintain an even playing field for participants. At a minimum, these laws govern matters like safeguarding private property against theft, protecting people from violence, enforcing legal contracts, preventing fraud, and collecting taxes. Conversely, even the most command-oriented economies operate using markets. How else would buying and selling occur? But the decisions of what will be produced and what prices will be charged are heavily regulated. Heavily regulated economies often have **underground economies**, which are markets where the buyers and sellers make transactions without the government’s approval.

The question of how to organize economic institutions is typically not a black-or-white choice between all market or all government, but instead involves a balancing act over the appropriate combination of market freedom and government rules.

Globalization



Cargo ships are one mode of transportation

for shipping goods in the global economy.
(Credit: Raul Valdez/Flickr Creative Commons)

The Rise of Globalization

Recent decades have seen a trend toward **globalization**, which is the expanding cultural, political, and economic connections between people around the world. One measure of this is the increased buying and selling of goods, services, and assets across national borders—in other words, international trade and financial capital flows.

Globalization has occurred for a number of reasons. Improvements in shipping, as illustrated by the container ship shown in [\[link\]](#), and air cargo have driven down transportation costs. Innovations in computing and telecommunications have made it easier and cheaper to manage long-distance economic connections of production and sales. Many valuable products and services in the modern economy can take the form of information—for example: computer software; financial advice; travel planning; music, books and movies; and blueprints for designing a building. These products and many others can be transported over telephones and computer networks at ever-lower costs. Finally, international agreements and treaties between countries have encouraged greater trade.

[\[link\]](#) presents one measure of globalization. It shows the percentage of domestic economic production that was exported for a selection of countries from 2010 to 2013, according to an entity known as The World Bank. **Exports** are the goods and services that are produced domestically and sold abroad. **Imports** are the goods and services that are produced abroad and then sold domestically. The size of total production in an economy is measured by the **gross domestic product (GDP)**. Thus, the ratio of exports divided by GDP measures what share of a country's total economic production is sold in other countries.

Country	2010	2011	2012	2013
Higher Income Countries				
United States	12.4	13.6	13.6	13.5
Belgium	76.2	81.4	82.2	82.8
Canada	29.1	30.7	30.0	30.1
France	26.0	27.8	28.1	28.3
Middle Income Countries				
Brazil	10.9	11.9	12.6	12.6
Mexico	29.9	31.2	32.6	31.7
South Korea	49.4	55.7	56.3	53.9
Lower Income Countries				
Chad	36.8	38.9	36.9	32.2
China	29.4	28.5	27.3	26.4
India	22.0	23.9	24.0	24.8
Nigeria	25.3	31.3	31.4	18.0

The Extent of Globalization (exports/GDP)(Source: <http://databank.worldbank.org/data/>)

In recent decades, the export/GDP ratio has generally risen, both worldwide and for the U.S. economy. Interestingly, the share of U.S. exports in proportion to the U.S. economy is well below the global average, in part because large economies like the United States can contain more of the

division of labor inside their national borders. However, smaller economies like Belgium, Korea, and Canada need to trade across their borders with other countries to take full advantage of division of labor, specialization, and economies of scale. In this sense, the enormous U.S. economy is less affected by globalization than most other countries.

[\[link\]](#) also shows that many medium and low income countries around the world, like Mexico and China, have also experienced a surge of globalization in recent decades. If an astronaut in orbit could put on special glasses that make all economic transactions visible as brightly colored lines and look down at Earth, the astronaut would see the planet covered with connections.

So, hopefully, you now have an idea of what economics is about. Before you move to any other chapter of study, be sure to read the very important appendix to this chapter called [The Use of Mathematics in Principles of Economics](#). It is essential that you learn more about how to read and use models in economics.

Note:**Decisions ... Decisions in the Social Media Age**

The world we live in today provides nearly instant access to a wealth of information. Consider that as recently as the late 1970s, the Farmer's Almanac, along with the Weather Bureau of the U.S. Department of Agriculture, were the primary sources American farmers used to determine when to plant and harvest their crops. Today, farmers are more likely to access, online, weather forecasts from the National Oceanic and Atmospheric Administration or watch the Weather Channel. After all, knowing the upcoming forecast could drive when to harvest crops. Consequently, knowing the upcoming weather could change the amount of crop harvested.

Some relatively new information forums, such as Facebook, are rapidly changing how information is distributed; hence, influencing decision making. In 2014, the Pew Research Center reported that 71% of online adults use Facebook. Facebook post topics range from the National Basketball Association, to celebrity singers and performers, to farmers.

Information helps us make decisions. Decisions as simple as what to wear today to how many reporters should be sent to cover a crash. Each of these decisions is an economic decision. After all, resources are scarce. If ten reporters are sent to cover an accident, they are not available to cover other stories or complete other tasks. Information provides the knowledge needed to make the best possible decisions on how to utilize scarce resources. Welcome to the world of economics!

Key Concepts and Summary

Societies can be organized as traditional, command, or market-oriented economies. Most societies are a mix. The last few decades have seen globalization evolve as a result of growth in commercial and financial networks that cross national borders, making businesses and workers from different economies increasingly interdependent.

Self-Check Questions

Exercise:

Problem:

The chapter defines *private enterprise* as a characteristic of market-oriented economies. What would *public enterprise* be? *Hint*: It is a characteristic of command economies.

Solution:

Public enterprise means the factors of production (resources and businesses) are owned and operated by the government.

Exercise:

Problem:

Why might Belgium, France, Italy, and Sweden have a higher export to GDP ratio than the United States?

Solution:

The United States is a large country economically speaking, so it has less need to trade internationally than the other countries mentioned. (This is the same reason that France and Italy have lower ratios than Belgium or Sweden.) One additional reason is that each of the other countries is a member of the European Union, where trade between members occurs without barriers to trade, like tariffs and quotas.

Review Questions**Exercise:****Problem:**

What are the three ways that societies can organize themselves economically?

Exercise:**Problem:**

What is globalization? How do you think it might have affected the economy over the past decade?

Critical Thinking Questions**Exercise:****Problem:**

Why do you think that most modern countries' economies are a mix of command and market types?

Exercise:

Problem:

Can you think of ways that globalization has helped you economically? Can you think of ways that it has not?

References

The Heritage Foundation. 2015. "2015 Index of Economic Freedom." Accessed March 11, 2015. <http://www.heritage.org/index/ranking>.

Garling, Caleb. "S.F. plane crash: Reporting, emotions on social media," *The San Francisco Chronicle*. July 7, 2013. <http://www.sfgate.com/news/article/S-F-plane-crash-Reporting-emotions-on-social-4651639.php>.

Irvine, Jessica. "Social Networking Sites are Factories of Modern Ideas." *The Sydney Morning Herald*. November 25, 2011. <http://www.smh.com.au/federal-politics/society-and-culture/social-networking-sites-are-factories-of-modern-ideas-20111124-1nwy3.html#ixzz2YZhPYeME>.

Pew Research Center. 2015. "Social Networking Fact Sheet." Accessed March 11, 2015. <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/>.

The World Bank Group. 2015. "World Data Bank." Accessed March 30, 2014. <http://databank.worldbank.org/data/>.

Glossary

command economy

an economy where economic decisions are passed down from government authority and where resources are owned by the government

exports

products (goods and services) made domestically and sold abroad

globalization

the trend in which buying and selling in markets have increasingly crossed national borders

gross domestic product (GDP)

measure of the size of total production in an economy

imports

products (goods and services) made abroad and then sold domestically

market

interaction between potential buyers and sellers; a combination of demand and supply

market economy

an economy where economic decisions are decentralized, resources are owned by private individuals, and businesses supply goods and services based on demand

private enterprise

system where the means of production (resources and businesses) are owned and operated by private individuals or groups of private individuals

traditional economy

typically an agricultural economy where things are done the same as they have always been done

underground economy

a market where the buyers and sellers make transactions in violation of one or more government regulations

Introduction to Choice in a World of Scarcity
class="introduction"
Choices and Tradeoffs

In general, the
higher the degree,
the higher the salary.
So why aren't more
people pursuing
higher degrees? The
short answer:
choices and
tradeoffs. (Credit:
modification of
work by "Jim, the
Photographer"/Flick
r Creative
Commons)



Note:**Choices ... To What Degree?**

In 2015, the median income for workers who hold master's degrees varies from males to females. The average of the two is \$2,951 weekly. Multiply this average by 52 weeks, and you get an average salary of \$153,452.

Compare that to the median weekly earnings for a full-time worker over 25 with no higher than a bachelor's degree: \$1,224 weekly and \$63,648 a year. What about those with no higher than a high school diploma in 2015? They earn just \$664 weekly and \$34,528 over 12 months. In other words, says the Bureau of Labor Statistics (BLS), earning a bachelor's degree boosted salaries 54% over what you would have earned if you had stopped your education after high school. A master's degree yields a salary almost double that of a high school diploma.

Given these statistics, we might expect a lot of people to choose to go to college and at least earn a bachelor's degree. Assuming that people want to improve their material well-being, it seems like they would make those choices that give them the greatest opportunity to consume goods and services. As it turns out, the analysis is not nearly as simple as this. In fact, in 2014, the BLS reported that while almost 88% of the population in the United States had a high school diploma, only 33.6% of 25–65 year olds had bachelor's degrees, and only 7.4% of 25–65 year olds in 2014 had earned a master's.

This brings us to the subject of this chapter: why people make the choices they make and how economists go about explaining those choices.

Note:**Introduction to Choice in a World of Scarcity**

In this chapter, you will learn about:

- How Individuals Make Choices Based on Their Budget Constraint
- The Production Possibilities Frontier and Social Choices
- Confronting Objections to the Economic Approach

You will learn quickly when you examine the relationship between economics and scarcity that choices involve tradeoffs. Every choice has a cost.

In 1968, the Rolling Stones recorded “You Can’t Always Get What You Want.” Economists chuckled, because they had been singing a similar tune for decades. English economist Lionel Robbins (1898–1984), in his *Essay on the Nature and Significance of Economic Science* in 1932, described not always getting what you want in this way:

"The time at our disposal is limited. There are only twenty-four hours in the day. We have to choose between the different uses to which they may be put. ... Everywhere we turn, if we choose one thing we must relinquish others which, in different circumstances, we would wish not to have relinquished. Scarcity of means to satisfy given ends is an almost ubiquitous condition of human nature."

Because people live in a world of scarcity, they cannot have all the time, money, possessions, and experiences they wish. Neither can society.

This chapter will continue our discussion of scarcity and the economic way of thinking by first introducing three critical concepts: opportunity cost, marginal decision making, and diminishing returns. Later, it will consider whether the economic way of thinking accurately describes either how choices *are* made or how they *should* be made.

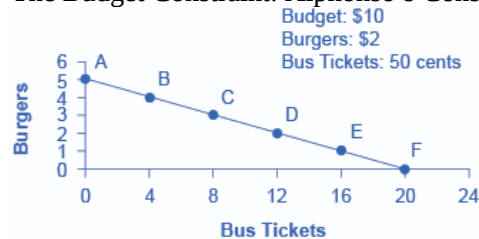
How Individuals Make Choices Based on Their Budget Constraint

By the end of this section, you will be able to:

- Calculate and graph budget constraints
- Explain opportunity sets and opportunity costs
- Evaluate the law of diminishing marginal utility
- Explain how marginal analysis and utility influence choices

Consider the typical consumer's budget problem. Consumers have a limited amount of income to spend on the things they need and want. Suppose Alphonso has \$10 in spending money each week that he can allocate between bus tickets for getting to work and the burgers that he eats for lunch. Burgers cost \$2 each, and bus tickets are 50 cents each. [\[link\]](#) shows Alphonso's **budget constraint**, that is, the outer boundary of his **opportunity set**. The opportunity set identifies all the opportunities for spending within his budget. The budget constraint indicates all the combinations of burgers and bus tickets Alphonso can afford when he exhausts his budget, given the prices of the two goods. (There are actually many different kinds of budget constraints. You will learn more about them in the chapter on [Consumer Choices](#).)

The Budget Constraint: Alphonso's Consumption Choice Opportunity Frontier



Each point on the budget constraint represents a combination of burgers and bus tickets whose total cost adds up to Alphonso's budget of \$10. The slope of the budget constraint is determined by the relative price of burgers and bus tickets. All along the budget set, giving up one burger means gaining four bus tickets.

The vertical axis in the figure shows burger purchases and the horizontal axis shows bus ticket purchases. If Alphonso spends all his money on burgers, he can afford five per week. ($\$10 \text{ per week} / \$2 \text{ per burger} = 5 \text{ burgers per week}$.) But if he does this, he will not be able to afford any bus tickets. This choice (zero bus tickets and five burgers) is shown by point A in the figure. Alternatively, if Alphonso spends all his money on bus tickets, he can afford 20 per week. ($\$10 \text{ per week} / \$0.50 \text{ per bus ticket} = 20 \text{ bus tickets per week}$.) Then, however, he will not be able to afford any burgers. This alternative choice (20 bus tickets and zero burgers) is shown by point F.

If Alphonso is like most people, he will choose some combination that includes both bus tickets and burgers. That is, he will choose some combination on the budget constraint that connects points A and F. Every point on (or inside) the constraint shows a combination of burgers and bus tickets that Alphonso can afford. Any point outside the constraint is not affordable, because it would cost more money than Alphonso has in his budget.

The budget constraint clearly shows the tradeoff Alphonso faces in choosing between burgers and bus tickets. Suppose he is currently at point D, where he can afford 12 bus tickets and two burgers. What would it cost Alphonso for one more burger? It would be natural to answer \$2, but that's not the way economists think. Instead they ask, how many bus tickets would Alphonso have to give up to get one more burger, while staying within his budget? The answer is four bus tickets. That is the true cost to Alphonso of one more burger.

The Concept of Opportunity Cost

Economists use the term **opportunity cost** to indicate what must be given up to obtain something that is desired. The idea behind opportunity cost is that the cost of one item is the lost opportunity to do or consume something else; in short, opportunity cost is the value of the next best alternative. For Alphonso, the opportunity cost of a burger is the four bus tickets he would have to give up. He would decide whether or not to choose the burger depending on whether the value of the burger exceeds the value of the forgone alternative—in this case, bus tickets. Since people must choose, they inevitably face tradeoffs in which they have to give up things they desire to get other things they desire more.

Note:

View this [website](#) for an example of opportunity cost—paying someone else to wait in line for you.



A fundamental principle of economics is that every choice has an opportunity cost. If you sleep through your economics class (not recommended, by the way), the opportunity cost is the learning you miss from not attending class. If you spend your income on video games, you cannot spend it on movies. If you choose to marry one person, you give up the opportunity to marry anyone else. In short, opportunity cost is all around us and part of human existence.

The following Work It Out feature shows a step-by-step analysis of a budget constraint calculation. Read through it to understand another important concept—slope—that is further explained in the appendix [The Use of Mathematics in Principles of Economics](#).

Note:

Understanding Budget Constraints

Budget constraints are easy to understand if you apply a little math. The appendix [The Use of Mathematics in Principles of Economics](#) explains all the math you are likely to need in this book. So if math is not your strength, you might want to take a look at the appendix.

Step 1: The equation for any budget constraint is:

Equation:

$$\text{Budget} = P_1 \times Q_1 + P_2 \times Q_2$$

where P and Q are the price and quantity of items purchased and Budget is the amount of income one has to spend.

Step 2. Apply the budget constraint equation to the scenario. In Alphonso's case, this works out to be:

Equation:

$$\text{Budget} = P_1 \times Q_1 + P_2 \times Q_2$$

\$10 budget = \$2 per burger \times quantity of burgers + \$0.50 per bus ticket \times quantity of bus tickets

$$\$10 = \$2 \times Q_{\text{burgers}} + \$0.50 \times Q_{\text{bus tickets}}$$

Step 3. Using a little algebra, we can turn this into the familiar equation of a line:

Equation:

$$y = b + mx$$

For Alphonso, this is:

Equation:

$$\$10 = \$2 \times Q_{\text{burgers}} + \$0.50 \times Q_{\text{bus tickets}}$$

Step 4. Simplify the equation. Begin by multiplying both sides of the equation by 2:

Equation:

$$2 \times 10 = 2 \times 2 \times Q_{\text{burgers}} + 2 \times 0.5 \times Q_{\text{bus tickets}}$$

$$20 = 4 \times Q_{\text{burgers}} + 1 \times Q_{\text{bus tickets}}$$

Step 5. Subtract one bus ticket from both sides:

Equation:

$$20 - Q_{\text{bus tickets}} = 4 \times Q_{\text{burgers}}$$

Divide each side by 4 to yield the answer:

Equation:

$$5 - 0.25 \times Q_{\text{bus tickets}} = Q_{\text{burgers}}$$

or

$$Q_{\text{burgers}} = 5 - 0.25 \times Q_{\text{bus tickets}}$$

Step 6. Notice that this equation fits the budget constraint in [\[link\]](#). The vertical intercept is 5 and the slope is -0.25 , just as the equation says. If you plug 20 bus tickets into the equation, you get 0 burgers. If you plug other numbers of bus tickets into the equation, you get the results shown in [\[link\]](#), which are the points on Alphonso's budget constraint.

Point	Quantity of Burgers (at \$2)	Quantity of Bus Tickets (at 50 cents)
A	5	0
B	4	4

Point	Quantity of Burgers (at \$2)	Quantity of Bus Tickets (at 50 cents)
C	3	8
D	2	12
E	1	16
F	0	20

Step 7. Notice that the slope of a budget constraint always shows the opportunity cost of the good which is on the horizontal axis. For Alphonso, the slope is -0.25 , indicating that for every four bus tickets he buys, Alphonso must give up 1 burger.

There are two important observations here. First, the algebraic sign of the slope is negative, which means that the only way to get more of one good is to give up some of the other. Second, the slope is defined as the price of bus tickets (whatever is on the horizontal axis in the graph) divided by the price of burgers (whatever is on the vertical axis), in this case $\$0.50/\$2 = 0.25$. So if you want to determine the opportunity cost quickly, just divide the two prices.

Identifying Opportunity Cost

In many cases, it is reasonable to refer to the opportunity cost as the price. If your cousin buys a new bicycle for \$300, then \$300 measures the amount of “other consumption” that he has given up. For practical purposes, there may be no special need to identify the specific alternative product or products that could have been bought with that \$300, but sometimes the price as measured in dollars may not accurately capture the true opportunity cost. This problem can loom especially large when costs of time are involved.

For example, consider a boss who decides that all employees will attend a two-day retreat to “build team spirit.” The out-of-pocket monetary cost of the event may involve hiring an outside consulting firm to run the retreat, as well as room and board for all participants. But an opportunity cost exists as well: during the two days of the retreat, none of the employees are doing any other work.

Attending college is another case where the opportunity cost exceeds the monetary cost. The out-of-pocket costs of attending college include tuition, books, room and board, and other expenses. But in addition, during the hours that you are attending class and studying, it is impossible to work at a paying job. Thus, college imposes both an out-of-pocket cost and an opportunity cost of lost earnings.

Note:

What is the opportunity cost associated with increased airport security measures?

After the terrorist plane hijackings on September 11, 2001, many steps were proposed to improve air travel safety. For example, the federal government could provide armed “sky marshals” who would travel inconspicuously with the rest of the passengers. The cost of having a sky marshal on every flight would be roughly \$3 billion per year. Retrofitting all U.S. planes with reinforced cockpit doors to make it harder for terrorists to take over the plane would have a price tag of \$450 million. Buying more sophisticated security equipment for airports, like three-dimensional baggage scanners and cameras linked to face recognition software, could cost another \$2 billion.

But the single biggest cost of greater airline security does not involve spending money. It is the opportunity cost of additional waiting time at the airport. According to the United States Department of Transportation

(DOT), more than 800 million passengers took plane trips in the United States in 2012. Since the 9/11 hijackings, security screening has become more intensive, and consequently, the procedure takes longer than in the past. Say that, on average, each air passenger spends an extra 30 minutes in the airport per trip. Economists commonly place a value on time to convert an opportunity cost in time into a monetary figure. Because many air travelers are relatively high-paid business people, conservative estimates set the average price of time for air travelers at \$20 per hour. By these back-of-the-envelope calculations, the opportunity cost of delays in airports could be as much as $800 \text{ million} \times 0.5 \text{ hours} \times \$20/\text{hour}$, or \$8 billion per year. Clearly, the opportunity costs of waiting time can be just as important as costs that involve direct spending.

In some cases, realizing the opportunity cost can alter behavior. Imagine, for example, that you spend \$8 on lunch every day at work. You may know perfectly well that bringing a lunch from home would cost only \$3 a day, so the opportunity cost of buying lunch at the restaurant is \$5 each day (that is, the \$8 buying lunch costs minus the \$3 your lunch from home would cost). \$5 each day does not seem to be that much. However, if you project what that adds up to in a year—250 days a year \times \$5 per day equals \$1,250, the cost, perhaps, of a decent vacation. If the opportunity cost is described as “a nice vacation” instead of “\$5 a day,” you might make different choices.

Marginal Decision-Making and Diminishing Marginal Utility

The budget constraint framework helps to emphasize that most choices in the real world are not about getting all of one thing or all of another; that is, they are not about choosing either the point at one end of the budget constraint or else the point all the way at the other end. Instead, most choices involve **marginal analysis**, which means comparing the benefits and costs of choosing a little more or a little less of a good.

People desire goods and services for the satisfaction or **utility** those goods and services provide. Utility, as we will see in the chapter on [Consumer Choices](#), is subjective but that does not make it less real. Economists typically assume that the more of some good one consumes (for example, slices of pizza), the more utility one obtains. At the same time, the utility a person receives from consuming the first unit of a good is typically more than the utility received from consuming the fifth or the tenth unit of that same good. When Alphonso chooses between burgers and bus tickets, for example, the first few bus rides that he chooses might provide him with a great deal of utility—perhaps they help him get to a job interview or a doctor’s appointment. But later bus rides might provide much less utility—they may only serve to kill time on a rainy day. Similarly, the first burger that Alphonso chooses to buy may be on a day when he missed breakfast and is ravenously hungry. However, if Alphonso has a burger every single day, the last few burgers may taste pretty boring. The general pattern that consumption of the first few units of any good tends to bring a higher level of utility to a person than consumption of later units is a common pattern. Economists refer to this pattern as the **law of diminishing marginal utility**, which means that as a person receives more of a good, the additional (or marginal) utility from each additional unit of the good declines. In other words, the first slice of pizza brings more satisfaction than the sixth.

The law of diminishing marginal utility explains why people and societies rarely make all-or-nothing choices. You would not say, “My favorite food is ice cream, so I will eat nothing but ice cream from now on.” Instead, even if you get a very high level of utility from your favorite food, if you ate it exclusively, the additional or marginal utility from those last few servings would not be very high. Similarly, most workers do not say: “I enjoy leisure, so I’ll never work.” Instead, workers recognize that even though some leisure is very nice, a combination of all leisure and no income is not so attractive. The budget constraint framework suggests that when people make choices in a world of scarcity, they will use marginal analysis and think about whether they would prefer a little more or a little less.

Sunk Costs

In the budget constraint framework, all decisions involve what will happen next: that is, what quantities of goods will you consume, how many hours will you work, or how much will you save. These decisions do not look back to past choices. Thus, the budget constraint framework assumes that **sunk costs**, which are costs that were incurred in the past and cannot be recovered, should not affect the current decision.

Consider the case of Selena, who pays \$8 to see a movie, but after watching the film for 30 minutes, she knows that it is truly terrible. Should she stay and watch the rest of the movie because she paid for the ticket, or should she leave? The money she spent is a sunk cost, and unless the theater manager is feeling kindly, Selena will not get a refund. But staying in the movie still means paying an opportunity cost in time. Her choice is whether to spend the next 90 minutes suffering through a cinematic disaster or to do something—anything—else. The lesson of sunk costs is to forget about the money and time that is irretrievably gone and instead to focus on the marginal costs and benefits of current and future options.

For people and firms alike, dealing with sunk costs can be frustrating. It often means admitting an earlier error in judgment. Many firms, for example, find it hard to give up on a new product that is doing poorly because they spent so much money in creating and launching the product. But the lesson of sunk costs is to ignore them and make decisions based on what will happen in the future.

From a Model with Two Goods to One of Many Goods

The budget constraint diagram containing just two goods, like most models used in this book, is not realistic. After all, in a modern economy people choose from thousands of goods. However, thinking about a model with many goods is a straightforward extension of what we discussed here. Instead of drawing just one budget constraint, showing the tradeoff between two goods, you can draw multiple budget constraints, showing the possible tradeoffs between many different pairs of goods. Or in more advanced classes in economics, you would use mathematical equations that include many possible goods and services that can be purchased, together with their quantities and prices, and show how the total spending on all goods and services is limited to the overall budget available. The graph with two goods that was presented here clearly illustrates that every choice has an opportunity cost, which is the point that does carry over to the real world.

Key Concepts and Summary

Economists see the real world as one of scarcity: that is, a world in which people's desires exceed what is possible. As a result, economic behavior involves tradeoffs in which individuals, firms, and society must give up something that they desire to obtain things that they desire more. Individuals face the tradeoff of what quantities of goods and services to consume. The budget constraint, which is the frontier of the opportunity set, illustrates the range of choices available. The slope of the budget constraint is determined by the relative price of the choices. Choices beyond the budget constraint are not affordable.

Opportunity cost measures cost by what is given up in exchange. Sometimes opportunity cost can be measured in money, but it is often useful to consider time as well, or to measure it in terms of the actual resources that must be given up.

Most economic decisions and tradeoffs are not all-or-nothing. Instead, they involve marginal analysis, which means they are about decisions on the margin, involving a little more or a little less. The law of diminishing marginal utility points out that as a person receives more of something—whether it is a specific good or another resource—the additional marginal gains tend to become smaller. Because sunk costs occurred in the past and cannot be recovered, they should be disregarded in making current decisions.

Self-Check Questions

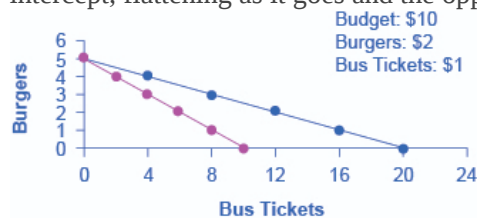
Exercise:

Problem:

Suppose Alphonso's town raised the price of bus tickets to \$1 per trip (while the price of burgers stayed at \$2 and his budget remained \$10 per week.) Draw Alphonso's new budget constraint. What happens to the opportunity cost of bus tickets?

Solution:

The opportunity cost of bus tickets is the number of burgers that must be given up to obtain one more bus ticket. Originally, when the price of bus tickets was 50 cents per trip, this opportunity cost was $0.50/2 = .25$ burgers. The reason for this is that at the original prices, one burger (\$2) costs the same as four bus tickets (\$0.50), so the opportunity cost of a burger is four bus tickets, and the opportunity cost of a bus ticket is .25 (the inverse of the opportunity cost of a burger). With the new, higher price of bus tickets, the opportunity cost rises to $1/2$ or 0.50. You can see this graphically since the slope of the new budget constraint is flatter than the original one. If Alphonso spends all of his budget on burgers, the higher price of bus tickets has no impact so the horizontal intercept of the budget constraint is the same. If he spends all of his budget on bus tickets, he can now afford only half as many, so the vertical intercept is half as much. In short, the budget constraint rotates clockwise around the horizontal intercept, flattening as it goes and the opportunity cost of bus tickets increases.



Review Questions

Exercise:

Problem: Explain why scarcity leads to tradeoffs.

Exercise:

Problem:

Explain why individuals make choices that are directly on the budget constraint, rather than inside the budget constraint or outside it.

Critical Thinking Question

Exercise:

Problem:

Suppose Alphonso's town raises the price of bus tickets from \$0.50 to \$1 and the price of burgers rises from \$2 to \$4. Why is the opportunity cost of bus tickets unchanged? Suppose Alphonso's weekly spending money increases from \$10 to \$20. How is his budget constraint affected from all three changes? Explain.

Problems

Use this information to answer the following 4 questions: Marie has a weekly budget of \$24, which she likes to spend on magazines and pies.

Exercise:

Problem:

If the price of a magazine is \$4 each, what is the maximum number of magazines she could buy in a week?

Exercise:

Problem: If the price of a pie is \$12, what is the maximum number of pies she could buy in a week?

Exercise:

Problem:

Draw Marie's budget constraint with pies on the horizontal axis and magazines on the vertical axis. What is the slope of the budget constraint?

Exercise:

Problem: What is Marie's opportunity cost of purchasing a pie?

References

Bureau of Labor Statistics, U.S. Department of Labor. 2015. "Median Weekly Earnings by Educational Attainment in 2014." Accessed March 27, 2015. <http://www.bls.gov/opub/ted/2015/median-weekly-earnings-by-education-gender-race-and-ethnicity-in-2014.htm>.

Robbins, Lionel. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan. 1932.

United States Department of Transportation. "Total Passengers on U.S Airlines and Foreign Airlines U.S. Flights Increased 1.3% in 2012 from 2011." Accessed October 2013. http://www.rita.dot.gov/bts/press_releases/bts016_13

Glossary

budget constraint

all possible consumption combinations of goods that someone can afford, given the prices of goods, when all income is spent; the boundary of the opportunity set

law of diminishing marginal utility

as we consume more of a good or service, the utility we get from additional units of the good or service tend to become smaller than what we received from earlier units

marginal analysis

examination of decisions on the margin, meaning a little more or a little less from the status quo

opportunity cost

measures cost by what is given up in exchange; opportunity cost measures the value of the forgone alternative

opportunity set

all possible combinations of consumption that someone can afford given the prices of goods and the individual's income

sunk costs

costs that are made in the past and cannot be recovered

utility

satisfaction, usefulness, or value one obtains from consuming goods and services

The Production Possibilities Frontier and Social Choices

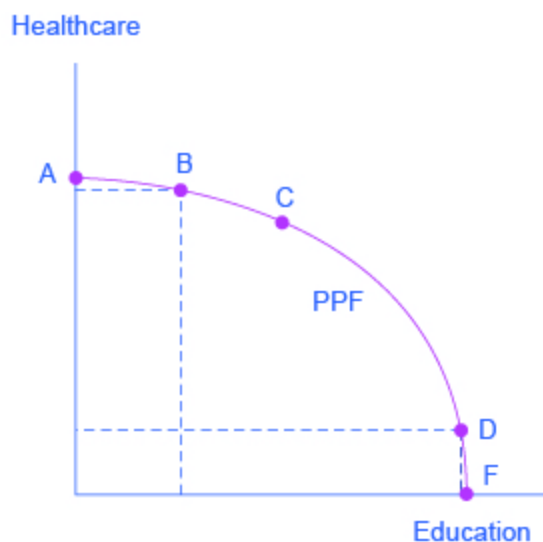
By the end of this section, you will be able to:

- Interpret production possibilities frontier graphs
- Contrast a budget constraint and a production possibilities frontier
- Explain the relationship between a production possibilities frontier and the law of diminishing returns
- Contrast productive efficiency and allocative efficiency
- Define comparative advantage

Just as individuals cannot have everything they want and must instead make choices, society as a whole cannot have everything it might want, either. This section of the chapter will explain the constraints faced by society, using a model called the **production possibilities frontier (PPF)**. There are more similarities than differences between individual choice and social choice. As you read this section, focus on the similarities.

Because society has limited resources (e.g., labor, land, capital, raw materials) at any point in time, there is a limit to the quantities of goods and services it can produce. Suppose a society desires two products, healthcare and education. This situation is illustrated by the production possibilities frontier in [\[link\]](#).

A Healthcare vs. Education Production Possibilities Frontier



This production possibilities frontier shows a tradeoff between devoting social resources to healthcare and devoting them to education. At A all resources go to healthcare and at B, most go to healthcare. At D most resources go to education, and at F, all go to education.

In [\[link\]](#), healthcare is shown on the vertical axis and education is shown on the horizontal axis. If the society were to allocate all of its resources to healthcare, it could produce at point A. But it would not have any resources to produce education. If it were to allocate all of its resources to education, it could produce at point F. Alternatively, the society could choose to produce any combination of healthcare and education shown on the production possibilities frontier. In effect, the production possibilities frontier plays the same role for society as the budget constraint plays for Alphonso. Society can choose any combination of the two goods on or inside the PPF. But it does not have enough resources to produce outside the PPF.

Most important, the production possibilities frontier clearly shows the tradeoff between healthcare and education. Suppose society has chosen to operate at point B, and it is considering producing more education. Because the PPF is downward sloping from left to right, the only way society can obtain more education is by giving up some healthcare. That is the tradeoff society faces. Suppose it considers moving from point B to point C. What would the opportunity cost be for the additional education? The opportunity cost would be the healthcare society has to give up. Just as with Alphonso's budget constraint, the opportunity cost is shown by the slope of the production possibilities frontier. By now you might be saying, "Hey, this PPF is sounding like the budget constraint." If so, read the following Clear It Up feature.

Note:

What's the difference between a budget constraint and a PPF?

There are two major differences between a budget constraint and a production possibilities frontier. The first is the fact that the budget constraint is a straight line. This is because its slope is given by the relative prices of the two goods. In contrast, the PPF has a curved shape because of the law of the diminishing returns. The second is the absence of specific numbers on the axes of the PPF. There are no specific numbers because we do not know the exact amount of resources this imaginary economy has, nor do we know how many resources it takes to produce healthcare and how many resources it takes to produce education. If this were a real world example, that data would be available. An additional reason for the lack of numbers is that there is no single way to measure levels of education and healthcare. However, when you think of improvements in education, you can think of accomplishments like more years of school completed, fewer high-school dropouts, and higher scores on standardized tests. When you think of improvements in healthcare, you can think of longer life expectancies, lower levels of infant mortality, and fewer outbreaks of disease.

Whether or not we have specific numbers, conceptually we can measure the opportunity cost of additional education as society moves from point B to point C on the PPF. The additional education is measured by the horizontal distance between B and C. The foregone healthcare is given by the vertical distance between B and C. The slope of the PPF between B and C is (approximately) the vertical distance (the “rise”) over the horizontal distance (the “run”). This is the opportunity cost of the additional education.

The Shape of the PPF and the Law of Diminishing Returns

The budget constraints presented earlier in this chapter, showing individual choices about what quantities of goods to consume, were all straight lines. The reason for these straight lines was that the slope of the budget constraint was determined by relative prices of the two goods in the consumption budget constraint. However, the production possibilities

frontier for healthcare and education was drawn as a curved line. Why does the PPF have a different shape?

To understand why the PPF is curved, start by considering point A at the top left-hand side of the PPF. At point A, all available resources are devoted to healthcare and none are left for education. This situation would be extreme and even ridiculous. For example, children are seeing a doctor every day, whether they are sick or not, but not attending school. People are having cosmetic surgery on every part of their bodies, but no high school or college education exists. Now imagine that some of these resources are diverted from healthcare to education, so that the economy is at point B instead of point A. Diverting some resources away from A to B causes relatively little reduction in health because the last few marginal dollars going into healthcare services are not producing much additional gain in health. However, putting those marginal dollars into education, which is completely without resources at point A, can produce relatively large gains. For this reason, the shape of the PPF from A to B is relatively flat, representing a relatively small drop-off in health and a relatively large gain in education.

Now consider the other end, at the lower right, of the production possibilities frontier. Imagine that society starts at choice D, which is devoting nearly all resources to education and very few to healthcare, and moves to point F, which is devoting *all* spending to education and none to healthcare. For the sake of concreteness, you can imagine that in the movement from D to F, the last few doctors must become high school science teachers, the last few nurses must become school librarians rather than dispensers of vaccinations, and the last few emergency rooms are turned into kindergartens. The gains to education from adding these last few resources to education are very small. However, the opportunity cost lost to health will be fairly large, and thus the slope of the PPF between D and F is steep, showing a large drop in health for only a small gain in education.

The lesson is not that society is likely to make an extreme choice like devoting no resources to education at point A or no resources to health at point F. Instead, the lesson is that the gains from committing additional marginal resources to education depend on how much is already being

spent. If on the one hand, very few resources are currently committed to education, then an increase in resources used can bring relatively large gains. On the other hand, if a large number of resources are already committed to education, then committing additional resources will bring relatively smaller gains.

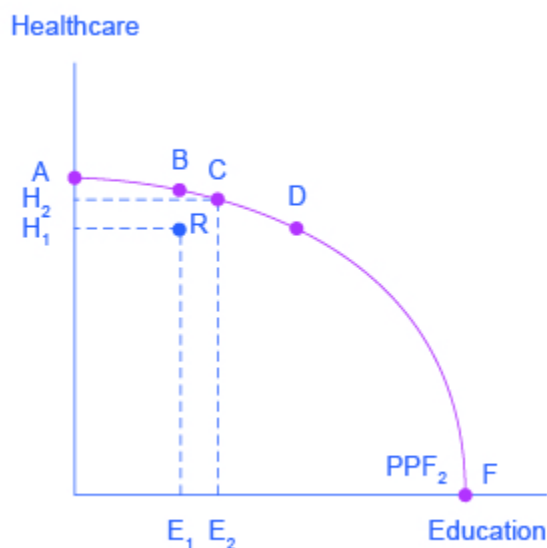
This pattern is common enough that it has been given a name: the **law of diminishing returns**, which holds that as additional increments of resources are added to a certain purpose, the marginal benefit from those additional increments will decline. When government spends a certain amount more on reducing crime, for example, the original gains in reducing crime could be relatively large. But additional increases typically cause relatively smaller reductions in crime, and paying for enough police and security to reduce crime to nothing at all would be tremendously expensive.

The curvature of the production possibilities frontier shows that as additional resources are added to education, moving from left to right along the horizontal axis, the original gains are fairly large, but gradually diminish. Similarly, as additional resources are added to healthcare, moving from bottom to top on the vertical axis, the original gains are fairly large, but again gradually diminish. In this way, the law of diminishing returns produces the outward-bending shape of the production possibilities frontier.

Productive Efficiency and Allocative Efficiency

The study of economics does not presume to tell a society what choice it should make along its production possibilities frontier. In a market-oriented economy with a democratic government, the choice will involve a mixture of decisions by individuals, firms, and government. However, economics can point out that some choices are unambiguously better than others. This observation is based on the concept of efficiency. In everyday usage, efficiency refers to lack of waste. An inefficient machine operates at high cost, while an efficient machine operates at lower cost, because it is not wasting energy or materials. An inefficient organization operates with long delays and high costs, while an efficient organization meets schedules, is focused, and performs within budget.

The production possibilities frontier can illustrate two kinds of efficiency: productive efficiency and allocative efficiency. [\[link\]](#) illustrates these ideas using a production possibilities frontier between healthcare and education. Productive and Allocative Efficiency



Productive efficiency means it is impossible to produce more of one good without decreasing the quantity that is produced of another good. Thus, all choices along a given PPF like B, C, and D display productive efficiency, but R does not. Allocative efficiency means that the particular mix of goods being produced—that is, the specific choice along the production possibilities frontier—represents the allocation that society most desires.

Productive efficiency means that, given the available inputs and technology, it is impossible to produce more of one good without decreasing the quantity that is produced of another good. All choices on the

PPF in [\[link\]](#), including A, B, C, D, and F, display productive efficiency. As a firm moves from any one of these choices to any other, either healthcare increases and education decreases or vice versa. However, any choice inside the production possibilities frontier is productively inefficient and wasteful because it is possible to produce more of one good, the other good, or some combination of both goods.

For example, point R is productively inefficient because it is possible at choice C to have more of both goods: education on the horizontal axis is higher at point C than point R (E_2 is greater than E_1), and healthcare on the vertical axis is also higher at point C than point R (H_2 is greater than H_1).

The particular mix of goods and services being produced—that is, the specific combination of healthcare and education chosen along the production possibilities frontier—can be shown as a ray (line) from the origin to a specific point on the PPF. Output mixes that had more healthcare (and less education) would have a steeper ray, while those with more education (and less healthcare) would have a flatter ray.

Allocative efficiency means that the particular mix of goods a society produces represents the combination that society most desires. How to determine what a society desires can be a controversial question, and is usually discussed in political science, sociology, and philosophy classes as well as in economics. At its most basic, allocative efficiency means producers supply the quantity of each product that consumers demand. Only one of the productively efficient choices will be the allocatively efficient choice for society as a whole.

Why Society Must Choose

Every economy faces two situations in which it may be able to expand consumption of all goods. In the first case, a society may discover that it has been using its resources inefficiently, in which case by improving efficiency and producing on the production possibilities frontier, it can have more of all goods (or at least more of some and less of none). In the second case, as resources grow over a period of years (e.g., more labor and more capital), the economy grows. As it does, the production possibilities frontier

for a society will tend to shift outward and society will be able to afford more of all goods.

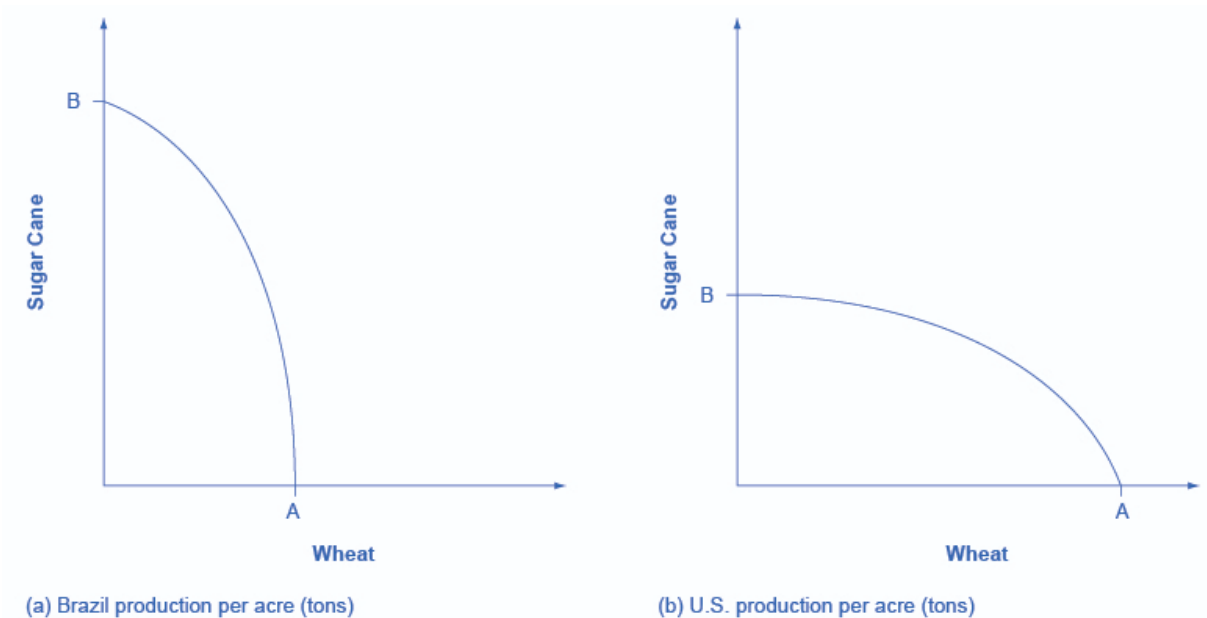
But improvements in productive efficiency take time to discover and implement, and economic growth happens only gradually. So, a society must choose between tradeoffs in the present. For government, this process often involves trying to identify where additional spending could do the most good and where reductions in spending would do the least harm. At the individual and firm level, the market economy coordinates a process in which firms seek to produce goods and services in the quantity, quality, and price that people want. But for both the government and the market economy in the short term, increases in production of one good typically mean offsetting decreases somewhere else in the economy.

The PPF and Comparative Advantage

While every society must choose how much of each good it should produce, it does not need to produce every single good it consumes. Often how much of a good a country decides to produce depends on how expensive it is to produce it versus buying it from a different country. As we saw earlier, the curvature of a country's PPF gives us information about the tradeoff between devoting resources to producing one good versus another. In particular, its slope gives the opportunity cost of producing one more unit of the good in the x-axis in terms of the other good (in the y-axis). Countries tend to have different opportunity costs of producing a specific good, either because of different climates, geography, technology or skills.

Suppose two countries, the US and Brazil, need to decide how much they will produce of two crops: sugar cane and wheat. Due to its climatic conditions, Brazil can produce a lot of sugar cane per acre but not much wheat. Conversely, the U.S. can produce a lot of wheat per acre, but not much sugar cane. Clearly, Brazil has a lower opportunity cost of producing sugar cane (in terms of wheat) than the U.S. The reverse is also true; the U.S. has a lower opportunity cost of producing wheat than Brazil. This can be illustrated by the PPFs of the two countries in [\[link\]](#)

Production Possibility Frontier for the U.S. and Brazil



The U.S. PPF is flatter than the Brazil PPF implying that the opportunity cost of wheat in term of sugar cane is lower in the U.S. than in Brazil. Conversely, the opportunity cost of sugar cane is lower in Brazil. The U.S. has comparative advantage in wheat and Brazil has comparative advantage in sugar cane.

When a country can produce a good at a lower opportunity cost than another country, we say that this country has a **comparative advantage** in that good. In our example, Brazil has a comparative advantage in sugar cane and the U.S. has a comparative advantage in wheat. One can easily see this with a simple observation of the extreme production points in the PPFs of the two countries. If Brazil devoted all of its resources to producing wheat, it would be producing at point A. If however it had devoted all of its resources to producing sugar cane instead, it would be producing a much larger amount, at point B. By moving from point A to point B Brazil would give up a relatively small quantity in wheat production to obtain a large production in sugar cane. The opposite is true for the U.S. If the U.S. moved from point A to B and produced only sugar cane, this would result in a large opportunity cost in terms of foregone wheat production.

The slope of the PPF gives the opportunity cost of producing an additional unit of wheat. While the slope is not constant throughout the PPFs, it is quite apparent that the PPF in Brazil is much steeper than in the U.S., and therefore the opportunity cost of wheat generally higher in Brazil. In the chapter on [International Trade](#) you will learn that countries' differences in comparative advantage determine which goods they will choose to produce and trade. When countries engage in trade, they specialize in the production of the goods that they have comparative advantage in, and trade part of that production for goods they do not have comparative advantage in. With trade, goods are produced where the opportunity cost is lowest, so total production increases, benefiting both trading parties.

Key Concepts and Summary

A production possibilities frontier defines the set of choices society faces for the combinations of goods and services it can produce given the resources available. The shape of the PPF is typically curved outward, rather than straight. Choices outside the PPF are unattainable and choices inside the PPF are wasteful. Over time, a growing economy will tend to shift the PPF outwards.

The law of diminishing returns holds that as increments of additional resources are devoted to producing something, the marginal increase in output will become smaller and smaller. All choices along a production possibilities frontier display productive efficiency; that is, it is impossible to use society's resources to produce more of one good without decreasing production of the other good. The specific choice along a production possibilities frontier that reflects the mix of goods society prefers is the choice with allocative efficiency. The curvature of the PPF is likely to differ by country, which results in different countries having comparative advantage in different goods. Total production can increase if countries specialize in the goods they have comparative advantage in and trade some of their production for the remaining goods.

Self-Check Questions

Exercise:

Problem:

Return to the example in [\[link\]](#). Suppose there is an improvement in medical technology that enables more healthcare to be provided with the same amount of resources. How would this affect the production possibilities curve and, in particular, how would it affect the opportunity cost of education?

Solution:

Because of the improvement in technology, the vertical intercept of the PPF would be at a higher level of healthcare. In other words, the PPF would rotate clockwise around the horizontal intercept. This would make the PPF steeper, corresponding to an increase in the opportunity cost of education, since resources devoted to education would now mean forgoing a greater quantity of healthcare.

Exercise:**Problem:**

Could a nation be producing in a way that is allocatively efficient, but productively inefficient?

Solution:

No. Allocative efficiency requires productive efficiency, because it pertains to choices along the production possibilities frontier.

Exercise:**Problem:**

What are the similarities between a consumer's budget constraint and society's production possibilities frontier, not just graphically but analytically?

Solution:

Both the budget constraint and the PPF show the constraint that each operates under. Both show a tradeoff between having more of one good but less of the other. Both show the opportunity cost graphically as the slope of the constraint (budget or PPF).

Review Questions

Exercise:

Problem: What is comparative advantage?

Exercise:

Problem: What does a production possibilities frontier illustrate?

Exercise:

Problem:

Why is a production possibilities frontier typically drawn as a curve, rather than a straight line?

Exercise:

Problem:

Explain why societies cannot make a choice above their production possibilities frontier and should not make a choice below it.

Exercise:

Problem: What are diminishing marginal returns?

Exercise:

Problem: What is productive efficiency? Allocative efficiency?

Critical Thinking Questions

Exercise:**Problem:**

During the Second World War, Germany's factories were decimated. It also suffered many human casualties, both soldiers and civilians. How did the war affect Germany's production possibilities curve?

Exercise:**Problem:**

It is clear that productive inefficiency is a waste since resources are being used in a way that produces less goods and services than a nation is capable of. Why is allocative inefficiency also wasteful?

Glossary

allocative efficiency

when the mix of goods being produced represents the mix that society most desires

comparative advantage

when a country can produce a good at a lower cost in terms of other goods; or, when a country has a lower opportunity cost of production

law of diminishing returns

as additional increments of resources are added to producing a good or service, the marginal benefit from those additional increments will decline

production possibilities frontier (PPF)

a diagram that shows the productively efficient combinations of two products that an economy can produce given the resources it has available.

productive efficiency

when it is impossible to produce more of one good (or service) without decreasing the quantity produced of another good (or service)

Introduction to Demand and Supply

class="introduction"

Farmer's Market

Organic vegetables and fruits that are grown and sold within a specific geographical region should, in theory, cost less than conventional produce because the transportation costs are less.

That is not, however, usually the case. (Credit: modification of work by Natalie Maynor/Flickr
r Creative Commons)

**Note:****Why Can We Not Get Enough of Organic?**

Organic food is increasingly popular, not just in the United States, but worldwide. At one time, consumers had to go to specialty stores or farmer's markets to find organic produce. Now it is available in most grocery stores. In short, organic is part of the mainstream.

Ever wonder why organic food costs more than conventional food? Why, say, does an organic Fuji apple cost \$1.99 a pound, while its conventional counterpart costs \$1.49 a pound? The same price relationship is true for just about every organic product on the market. If many organic foods are locally grown, would they not take less time to get to market and therefore be cheaper? What are the forces that keep those prices from coming down? Turns out those forces have a lot to do with this chapter's topic: demand and supply.

Note:**Introduction to Demand and Supply**

In this chapter, you will learn about:

- Demand, Supply, and Equilibrium in Markets for Goods and Services
- Shifts in Demand and Supply for Goods and Services
- Changes in Equilibrium Price and Quantity: The Four-Step Process
- Price Ceilings and Price Floors

An auction bidder pays thousands of dollars for a dress Whitney Houston wore. A collector spends a small fortune for a few drawings by John Lennon. People usually react to purchases like these in two ways: their jaw drops because they think these are high prices to pay for such goods or they think these are rare, desirable items and the amount paid seems right.

Note:

Visit this [website](#) to read a list of bizarre items that have been purchased for their ties to celebrities. These examples represent an interesting facet of demand and supply.



When economists talk about prices, they are less interested in making judgments than in gaining a practical understanding of what determines prices and why prices change. Consider a price most of us contend with weekly: that of a gallon of gas. Why was the average price of gasoline in the United States \$3.71 per gallon in June 2014? Why did the price for gasoline fall sharply to \$2.07 per gallon by January 2015? To explain these price movements, economists focus on the determinants of what gasoline buyers are willing to pay and what gasoline sellers are willing to accept.

As it turns out, the price of gasoline in June of any given year is nearly always higher than the price in January of that same year; over recent decades, gasoline prices in midsummer have averaged about 10 cents per gallon more than their midwinter low. The likely reason is that people drive more in the summer, and are also willing to pay more for gas, but that does not explain how steeply gas prices fell. Other factors were at work during those six months, such as increases in supply and decreases in the demand for crude oil.

This chapter introduces the economic model of demand and supply—one of the most powerful models in all of economics. The discussion here begins by examining how demand and supply determine the price and the quantity sold in markets for goods and services, and how changes in demand and supply lead to changes in prices and quantities.

Demand, Supply, and Equilibrium in Markets for Goods and Services

By the end of this section, you will be able to:

- Explain demand, quantity demanded, and the law of demand
- Identify a demand curve and a supply curve
- Explain supply, quantity supply, and the law of supply
- Explain equilibrium, equilibrium price, and equilibrium quantity

First let's first focus on what economists mean by demand, what they mean by supply, and then how demand and supply interact in a market.

Demand for Goods and Services

Economists use the term **demand** to refer to the amount of some good or service consumers are willing and able to purchase at each price. Demand is based on needs and wants—a consumer may be able to differentiate between a need and a want, but from an economist's perspective they are the same thing. Demand is also based on ability to pay. If you cannot pay for it, you have no effective demand.

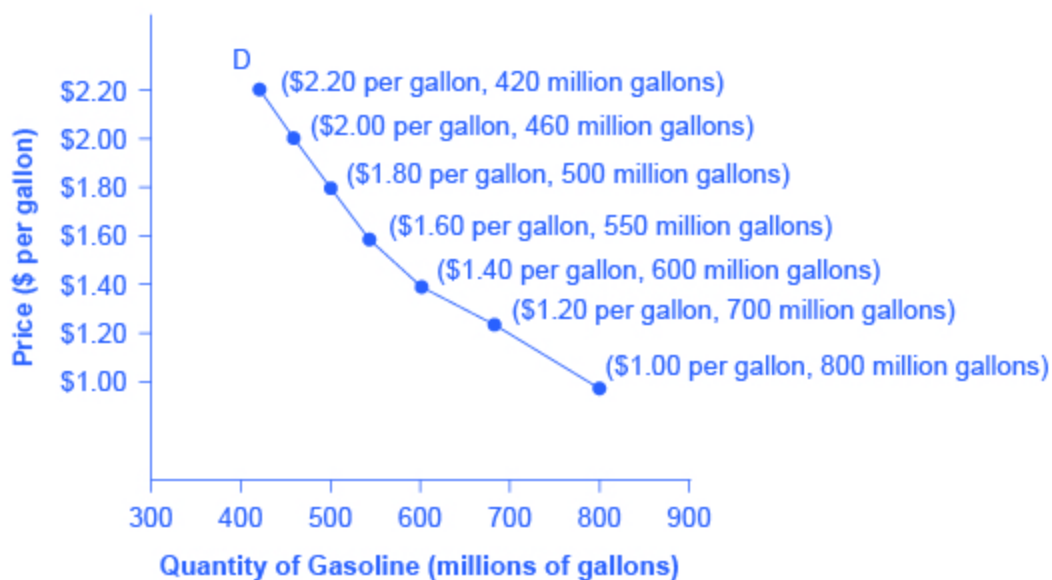
What a buyer pays for a unit of the specific good or service is called **price**. The total number of units purchased at that price is called the **quantity demanded**. A rise in price of a good or service almost always decreases the quantity demanded of that good or service. Conversely, a fall in price will increase the quantity demanded. When the price of a gallon of gasoline goes up, for example, people look for ways to reduce their consumption by combining several errands, commuting by carpool or mass transit, or taking weekend or vacation trips closer to home. Economists call this inverse relationship between price and quantity demanded the **law of demand**. The law of demand assumes that all other variables that affect demand (to be explained in the next module) are held constant.

An example from the market for gasoline can be shown in the form of a table or a graph. A table that shows the quantity demanded at each price, such as [\[link\]](#), is called a **demand schedule**. Price in this case is measured in dollars per gallon of gasoline. The quantity demanded is measured in

millions of gallons over some time period (for example, per day or per year) and over some geographic area (like a state or a country). A **demand curve** shows the relationship between price and quantity demanded on a graph like [\[link\]](#), with quantity on the horizontal axis and the price per gallon on the vertical axis. (Note that this is an exception to the normal rule in mathematics that the independent variable (x) goes on the horizontal axis and the dependent variable (y) goes on the vertical. Economics is not math.)

The demand schedule shown by [\[link\]](#) and the demand curve shown by the graph in [\[link\]](#) are two ways of describing the same relationship between price and quantity demanded.

A Demand Curve for Gasoline



The demand schedule shows that as price rises, quantity demanded decreases, and vice versa. These points are then graphed, and the line connecting them is the demand curve (D). The downward slope of the demand curve again illustrates the law of demand—the inverse relationship between prices and quantity demanded.

Price (per gallon)	Quantity Demanded (millions of gallons)
\$1.00	800
\$1.20	700
\$1.40	600
\$1.60	550
\$1.80	500
\$2.00	460
\$2.20	420

Price and Quantity Demanded of Gasoline

Demand curves will appear somewhat different for each product. They may appear relatively steep or flat, or they may be straight or curved. Nearly all demand curves share the fundamental similarity that they slope down from left to right. So demand curves embody the law of demand: As the price increases, the quantity demanded decreases, and conversely, as the price decreases, the quantity demanded increases.

Confused about these different types of demand? Read the next Clear It Up feature.

Note:

Is demand the same as quantity demanded?

In economic terminology, demand is not the same as quantity demanded. When economists talk about demand, they mean the relationship between a range of prices and the quantities demanded at those prices, as illustrated by a demand curve or a demand schedule. When economists talk about quantity demanded, they mean only a certain point on the demand curve, or

one quantity on the demand schedule. In short, demand refers to the curve and quantity demanded refers to the (specific) point on the curve.

Supply of Goods and Services

When economists talk about **supply**, they mean the amount of some good or service a producer is willing to supply at each price. Price is what the producer receives for selling one unit of a good or service. A rise in price almost always leads to an increase in the **quantity supplied** of that good or service, while a fall in price will decrease the quantity supplied. When the price of gasoline rises, for example, it encourages profit-seeking firms to take several actions: expand exploration for oil reserves; drill for more oil; invest in more pipelines and oil tankers to bring the oil to plants where it can be refined into gasoline; build new oil refineries; purchase additional pipelines and trucks to ship the gasoline to gas stations; and open more gas stations or keep existing gas stations open longer hours. Economists call this positive relationship between price and quantity supplied—that a higher price leads to a higher quantity supplied and a lower price leads to a lower quantity supplied—the **law of supply**. The law of supply assumes that all other variables that affect supply (to be explained in the next module) are held constant.

Still unsure about the different types of supply? See the following Clear It Up feature.

Note:

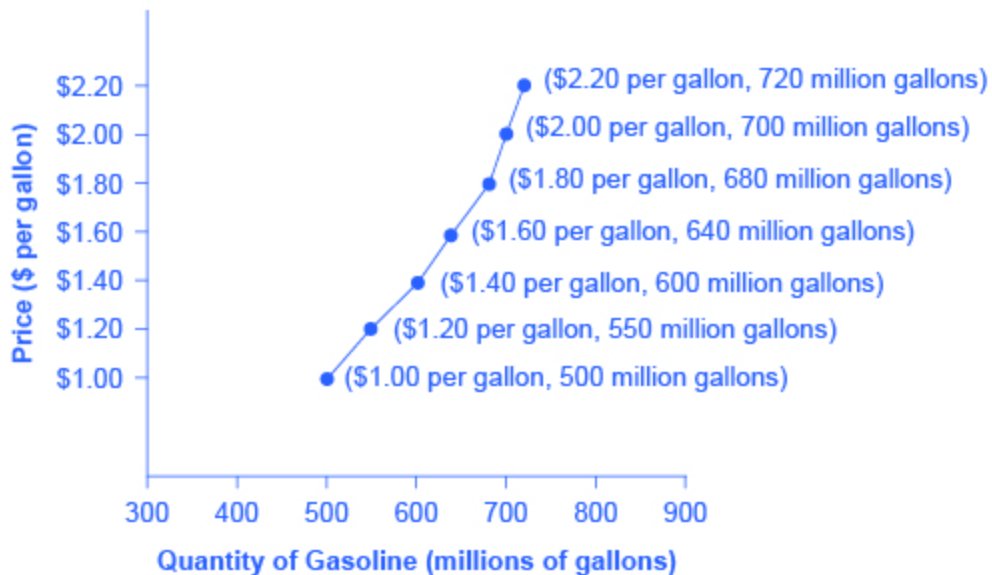
Is supply the same as quantity supplied?

In economic terminology, supply is not the same as quantity supplied. When economists refer to supply, they mean the relationship between a range of prices and the quantities supplied at those prices, a relationship that can be illustrated with a supply curve or a supply schedule. When economists refer to quantity supplied, they mean only a certain point on the supply curve, or one quantity on the supply schedule. In short, supply

refers to the curve and quantity supplied refers to the (specific) point on the curve.

[\[link\]](#) illustrates the law of supply, again using the market for gasoline as an example. Like demand, supply can be illustrated using a table or a graph. A **supply schedule** is a table, like [\[link\]](#), that shows the quantity supplied at a range of different prices. Again, price is measured in dollars per gallon of gasoline and quantity supplied is measured in millions of gallons. A **supply curve** is a graphic illustration of the relationship between price, shown on the vertical axis, and quantity, shown on the horizontal axis. The supply schedule and the supply curve are just two different ways of showing the same information. Notice that the horizontal and vertical axes on the graph for the supply curve are the same as for the demand curve.

A Supply Curve for Gasoline



The supply schedule is the table that shows quantity supplied of gasoline at each price. As price rises, quantity supplied also increases, and vice versa. The supply curve (S) is created by graphing the points from the supply schedule and then connecting them. The upward slope of the supply curve illustrates the law of supply—that a higher price leads to a higher quantity supplied, and vice versa.

Price (per gallon)	Quantity Supplied (millions of gallons)
\$1.00	500
\$1.20	550
\$1.40	600
\$1.60	640
\$1.80	680
\$2.00	700
\$2.20	720

Price and Supply of Gasoline

The shape of supply curves will vary somewhat according to the product: steeper, flatter, straighter, or curved. Nearly all supply curves, however, share a basic similarity: they slope up from left to right and illustrate the law of supply: as the price rises, say, from \$1.00 per gallon to \$2.20 per gallon, the quantity supplied increases from 500 gallons to 720 gallons. Conversely, as the price falls, the quantity supplied decreases.

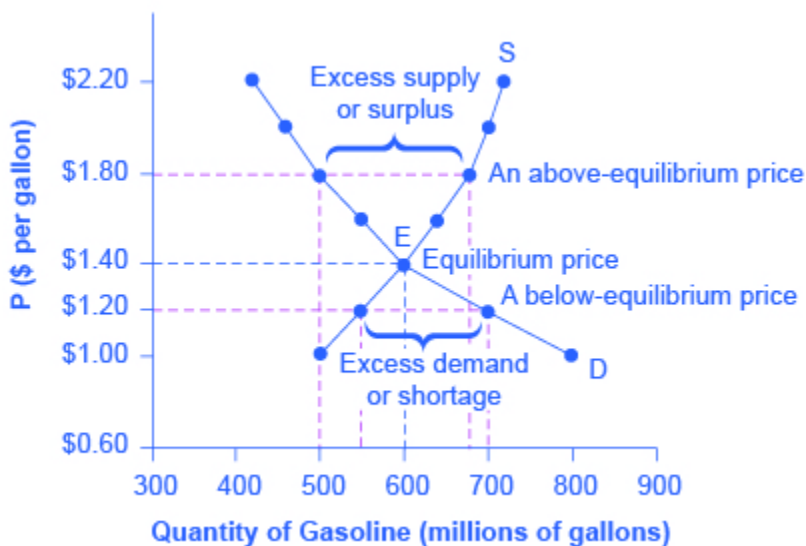
Equilibrium—Where Demand and Supply Intersect

Because the graphs for demand and supply curves both have price on the vertical axis and quantity on the horizontal axis, the demand curve and supply curve for a particular good or service can appear on the same graph.

Together, demand and supply determine the price and the quantity that will be bought and sold in a market.

[\[link\]](#) illustrates the interaction of demand and supply in the market for gasoline. The demand curve (D) is identical to [\[link\]](#). The supply curve (S) is identical to [\[link\]](#). [\[link\]](#) contains the same information in tabular form.

Demand and Supply for Gasoline



The demand curve (D) and the supply curve (S) intersect at the equilibrium point E, with a price of \$1.40 and a quantity of 600. The equilibrium is the only price where quantity demanded is equal to quantity supplied. At a price above equilibrium like \$1.80, quantity supplied exceeds the quantity demanded, so there is excess supply. At a price below equilibrium such as \$1.20, quantity demanded exceeds quantity supplied, so there is excess demand.

Price (per gallon)	Quantity demanded (millions of gallons)	Quantity supplied (millions of gallons)
\$1.00	800	500
\$1.20	700	550
\$1.40	600	600
\$1.60	550	640
\$1.80	500	680
\$2.00	460	700
\$2.20	420	720

Price, Quantity Demanded, and Quantity Supplied

Remember this: When two lines on a diagram cross, this intersection usually means something. The point where the supply curve (S) and the demand curve (D) cross, designated by point E in [\[link\]](#), is called the **equilibrium**. The **equilibrium price** is the only price where the plans of consumers and the plans of producers agree—that is, where the amount of the product consumers want to buy (quantity demanded) is equal to the amount producers want to sell (quantity supplied). This common quantity is called the **equilibrium quantity**. At any other price, the quantity demanded does not equal the quantity supplied, so the market is not in equilibrium at that price.

In [\[link\]](#), the equilibrium price is \$1.40 per gallon of gasoline and the equilibrium quantity is 600 million gallons. If you had only the demand and supply schedules, and not the graph, you could find the equilibrium by looking for the price level on the tables where the quantity demanded and the quantity supplied are equal.

The word “equilibrium” means “balance.” If a market is at its equilibrium price and quantity, then it has no reason to move away from that point. However, if a market is not at equilibrium, then economic pressures arise to move the market toward the equilibrium price and the equilibrium quantity.

Imagine, for example, that the price of a gallon of gasoline was above the equilibrium price—that is, instead of \$1.40 per gallon, the price is \$1.80 per gallon. This above-equilibrium price is illustrated by the dashed horizontal line at the price of \$1.80 in [\[link\]](#). At this higher price, the quantity demanded drops from 600 to 500. This decline in quantity reflects how consumers react to the higher price by finding ways to use less gasoline.

Moreover, at this higher price of \$1.80, the quantity of gasoline supplied rises from the 600 to 680, as the higher price makes it more profitable for gasoline producers to expand their output. Now, consider how quantity demanded and quantity supplied are related at this above-equilibrium price. Quantity demanded has fallen to 500 gallons, while quantity supplied has risen to 680 gallons. In fact, at any above-equilibrium price, the quantity supplied exceeds the quantity demanded. We call this an **excess supply** or a **surplus**.

With a surplus, gasoline accumulates at gas stations, in tanker trucks, in pipelines, and at oil refineries. This accumulation puts pressure on gasoline sellers. If a surplus remains unsold, those firms involved in making and selling gasoline are not receiving enough cash to pay their workers and to cover their expenses. In this situation, some producers and sellers will want to cut prices, because it is better to sell at a lower price than not to sell at all. Once some sellers start cutting prices, others will follow to avoid losing sales. These price reductions in turn will stimulate a higher quantity demanded. So, if the price is above the equilibrium level, incentives built into the structure of demand and supply will create pressures for the price to fall toward the equilibrium.

Now suppose that the price is below its equilibrium level at \$1.20 per gallon, as the dashed horizontal line at this price in [\[link\]](#) shows. At this lower price, the quantity demanded increases from 600 to 700 as drivers take longer trips, spend more minutes warming up the car in the driveway in wintertime, stop sharing rides to work, and buy larger cars that get fewer

miles to the gallon. However, the below-equilibrium price reduces gasoline producers' incentives to produce and sell gasoline, and the quantity supplied falls from 600 to 550.

When the price is below equilibrium, there is **excess demand**, or a **shortage**—that is, at the given price the quantity demanded, which has been stimulated by the lower price, now exceeds the quantity supplied, which had been depressed by the lower price. In this situation, eager gasoline buyers mob the gas stations, only to find many stations running short of fuel. Oil companies and gas stations recognize that they have an opportunity to make higher profits by selling what gasoline they have at a higher price. As a result, the price rises toward the equilibrium level. Read [Demand, Supply, and Efficiency](#) for more discussion on the importance of the demand and supply model.

Key Concepts and Summary

A demand schedule is a table that shows the quantity demanded at different prices in the market. A demand curve shows the relationship between quantity demanded and price in a given market on a graph. The law of demand states that a higher price typically leads to a lower quantity demanded.

A supply schedule is a table that shows the quantity supplied at different prices in the market. A supply curve shows the relationship between quantity supplied and price on a graph. The law of supply says that a higher price typically leads to a higher quantity supplied.

The equilibrium price and equilibrium quantity occur where the supply and demand curves cross. The equilibrium occurs where the quantity demanded is equal to the quantity supplied. If the price is below the equilibrium level, then the quantity demanded will exceed the quantity supplied. Excess demand or a shortage will exist. If the price is above the equilibrium level, then the quantity supplied will exceed the quantity demanded. Excess supply or a surplus will exist. In either case, economic pressures will push the price toward the equilibrium level.

Self-Check Question

Exercise:

Problem:

Review [\[link\]](#). Suppose the price of gasoline is \$1.60 per gallon. Is the quantity demanded higher or lower than at the equilibrium price of \$1.40 per gallon? And what about the quantity supplied? Is there a shortage or a surplus in the market? If so, of how much?

Solution:

Since \$1.60 per gallon is above the equilibrium price, the quantity demanded would be lower at 550 gallons and the quantity supplied would be higher at 640 gallons. (These results are due to the laws of demand and supply, respectively.) The outcome of lower Q_d and higher Q_s would be a surplus in the gasoline market of $640 - 550 = 90$ gallons.

Review Questions

Exercise:

Problem: What determines the level of prices in a market?

Exercise:

Problem:

What does a downward-sloping demand curve mean about how buyers in a market will react to a higher price?

Exercise:

Problem:

Will demand curves have the same exact shape in all markets? If not, how will they differ?

Exercise:

Problem:

Will supply curves have the same shape in all markets? If not, how will they differ?

Exercise:

Problem:

What is the relationship between quantity demanded and quantity supplied at equilibrium? What is the relationship when there is a shortage? What is the relationship when there is a surplus?

Exercise:

Problem:

How can you locate the equilibrium point on a demand and supply graph?

Exercise:

Problem:

If the price is above the equilibrium level, would you predict a surplus or a shortage? If the price is below the equilibrium level, would you predict a surplus or a shortage? Why?

Exercise:

Problem:

When the price is above the equilibrium, explain how market forces move the market price to equilibrium. Do the same when the price is below the equilibrium.

Exercise:

Problem:

What is the difference between the demand and the quantity demanded of a product, say milk? Explain in words and show the difference on a graph with a demand curve for milk.

Exercise:**Problem:**

What is the difference between the supply and the quantity supplied of a product, say milk? Explain in words and show the difference on a graph with the supply curve for milk.

Critical Thinking Questions**Exercise:****Problem:**

Review [\[link\]](#). Suppose the government decided that, since gasoline is a necessity, its price should be legally capped at \$1.30 per gallon. What do you anticipate would be the outcome in the gasoline market?

Exercise:**Problem:**

Explain why the following statement is false: “In the goods market, no buyer would be willing to pay more than the equilibrium price.”

Exercise:**Problem:**

Explain why the following statement is false: “In the goods market, no seller would be willing to sell for less than the equilibrium price.”

Problems

Exercise:

Problem:

Review [\[link\]](#) again. Suppose the price of gasoline is \$1.00. Will the quantity demanded be lower or higher than at the equilibrium price of \$1.40 per gallon? Will the quantity supplied be lower or higher? Is there a shortage or a surplus in the market? If so, of how much?

References

Costanza, Robert, and Lisa Wainger. "No Accounting For Nature: How Conventional Economics Distorts the Value of Things." *The Washington Post*. September 2, 1990.

European Commission: Agriculture and Rural Development. 2013. "Overview of the CAP Reform: 2014-2024." Accessed April 13, 2015. <http://ec.europa.eu/agriculture/cap-post-2013/>.

Radford, R. A. "The Economic Organisation of a P.O.W. Camp." *Economica*. no. 48 (1945): 189-201. <http://www.jstor.org/stable/2550133>.

Glossary

demand curve

a graphic representation of the relationship between price and quantity demanded of a certain good or service, with quantity on the horizontal axis and the price on the vertical axis

demand schedule

a table that shows a range of prices for a certain good or service and the quantity demanded at each price

demand

the relationship between price and the quantity demanded of a certain good or service

equilibrium price

the price where quantity demanded is equal to quantity supplied

equilibrium quantity

the quantity at which quantity demanded and quantity supplied are equal for a certain price level

equilibrium

the situation where quantity demanded is equal to the quantity supplied; the combination of price and quantity where there is no economic pressure from surpluses or shortages that would cause price or quantity to change

excess demand

at the existing price, the quantity demanded exceeds the quantity supplied; also called a shortage

excess supply

at the existing price, quantity supplied exceeds the quantity demanded; also called a surplus

law of demand

the common relationship that a higher price leads to a lower quantity demanded of a certain good or service and a lower price leads to a higher quantity demanded, while all other variables are held constant

law of supply

the common relationship that a higher price leads to a greater quantity supplied and a lower price leads to a lower quantity supplied, while all other variables are held constant

price

what a buyer pays for a unit of the specific good or service

quantity demanded

the total number of units of a good or service consumers are willing to purchase at a given price

quantity supplied

the total number of units of a good or service producers are willing to sell at a given price

shortage

at the existing price, the quantity demanded exceeds the quantity supplied; also called excess demand

supply curve

a line that shows the relationship between price and quantity supplied on a graph, with quantity supplied on the horizontal axis and price on the vertical axis

supply schedule

a table that shows a range of prices for a good or service and the quantity supplied at each price

supply

the relationship between price and the quantity supplied of a certain good or service

surplus

at the existing price, quantity supplied exceeds the quantity demanded; also called excess supply

Shifts in Demand and Supply for Goods and Services

By the end of this section, you will be able to:

- Identify factors that affect demand
- Graph demand curves and demand shifts
- Identify factors that affect supply
- Graph supply curves and supply shifts

The previous module explored how price affects the quantity demanded and the quantity supplied. The result was the demand curve and the supply curve. Price, however, is not the only thing that influences demand. Nor is it the only thing that influences supply. For example, how is demand for vegetarian food affected if, say, health concerns cause more consumers to avoid eating meat? Or how is the supply of diamonds affected if diamond producers discover several new diamond mines? What are the major factors, in addition to the price, that influence demand or supply?

Note:

Visit this [website](#) to read a brief note on how marketing strategies can influence supply and demand of products.



What Factors Affect Demand?

We defined demand as the amount of some product a consumer is willing and able to purchase at each price. That suggests at least two factors in addition to price that affect demand. Willingness to purchase suggests a

desire, based on what economists call tastes and preferences. If you neither need nor want something, you will not buy it. Ability to purchase suggests that income is important. Professors are usually able to afford better housing and transportation than students, because they have more income. Prices of related goods can affect demand also. If you need a new car, the price of a Honda may affect your demand for a Ford. Finally, the size or composition of the population can affect demand. The more children a family has, the greater their demand for clothing. The more driving-age children a family has, the greater their demand for car insurance, and the less for diapers and baby formula.

These factors matter both for demand by an individual and demand by the market as a whole. Exactly how do these various factors affect demand, and how do we show the effects graphically? To answer those questions, we need the *ceteris paribus* assumption.

The *Ceteris Paribus* Assumption

A demand curve or a supply curve is a relationship between two, and only two, variables: quantity on the horizontal axis and price on the vertical axis. The assumption behind a demand curve or a supply curve is that no relevant economic factors, other than the product's price, are changing. Economists call this assumption **ceteris paribus**, a Latin phrase meaning “other things being equal.” Any given demand or supply curve is based on the *ceteris paribus* assumption that all else is held equal. A demand curve or a supply curve is a relationship between two, and only two, variables when all other variables are kept constant. If all else is not held equal, then the laws of supply and demand will not necessarily hold, as the following Clear It Up feature shows.

Note:

When does *ceteris paribus* apply?

Ceteris paribus is typically applied when we look at how changes in price affect demand or supply, but *ceteris paribus* can be applied more generally. In the real world, demand and supply depend on more factors than just

price. For example, a consumer's demand depends on income and a producer's supply depends on the cost of producing the product. How can we analyze the effect on demand or supply if multiple factors are changing at the same time—say price rises and income falls? The answer is that we examine the changes one at a time, assuming the other factors are held constant.

For example, we can say that an increase in the price reduces the amount consumers will buy (assuming income, and anything else that affects demand, is unchanged). Additionally, a decrease in income reduces the amount consumers can afford to buy (assuming price, and anything else that affects demand, is unchanged). This is what the *ceteris paribus* assumption really means. In this particular case, after we analyze each factor separately, we can combine the results. The amount consumers buy falls for two reasons: first because of the higher price and second because of the lower income.

How Does Income Affect Demand?

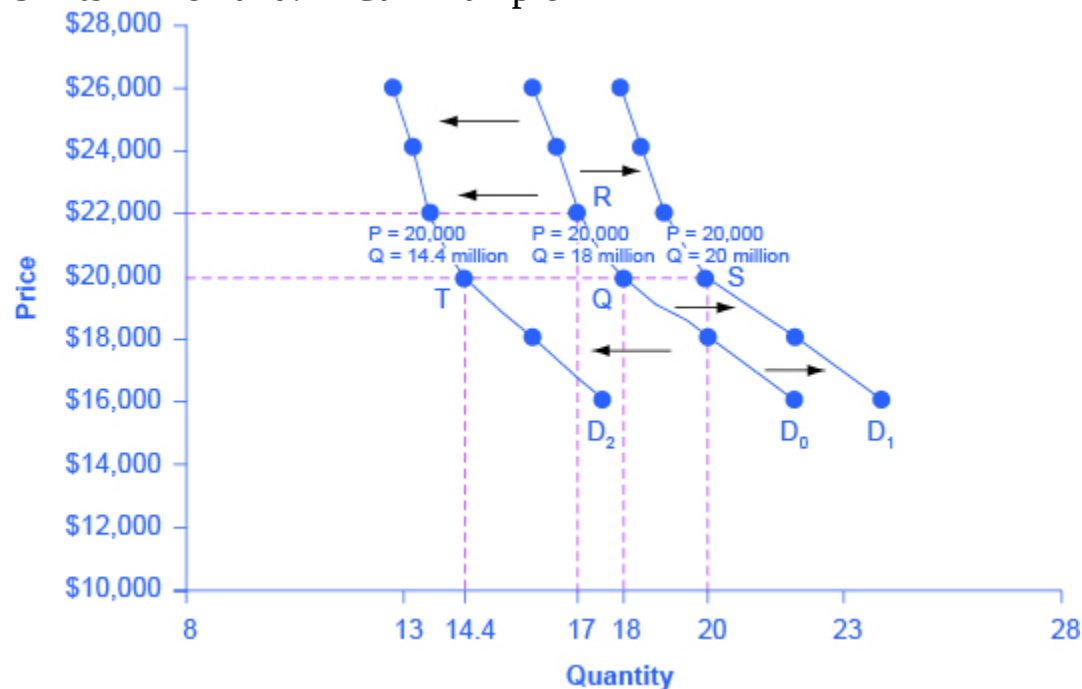
Let's use income as an example of how factors other than price affect demand. [\[link\]](#) shows the initial demand for automobiles as D_0 . At point Q, for example, if the price is \$20,000 per car, the quantity of cars demanded is 18 million. D_0 also shows how the quantity of cars demanded would change as a result of a higher or lower price. For example, if the price of a car rose to \$22,000, the quantity demanded would decrease to 17 million, at point R.

The original demand curve D_0 , like every demand curve, is based on the *ceteris paribus* assumption that no other economically relevant factors change. Now imagine that the economy expands in a way that raises the incomes of many people, making cars more affordable. How will this affect demand? How can we show this graphically?

Return to [\[link\]](#). The price of cars is still \$20,000, but with higher incomes, the quantity demanded has now increased to 20 million cars, shown at point S. As a result of the higher income levels, the demand curve shifts to the right to the new demand curve D_1 , indicating an increase in demand. [\[link\]](#)

shows clearly that this increased demand would occur at every price, not just the original one.

Shifts in Demand: A Car Example



Increased demand means that at every given price, the quantity demanded is higher, so that the demand curve shifts to the right from D_0 to D_1 . Decreased demand means that at every given price, the quantity demanded is lower, so that the demand curve shifts to the left from D_0 to D_2 .

Price	Decrease to D_2	Original Quantity Demanded D_0	Increase to D_1
\$16,000	17.6 million	22.0 million	24.0 million

Price	Decrease to D_2	Original Quantity Demanded D_0	Increase to D_1
\$18,000	16.0 million	20.0 million	22.0 million
\$20,000	14.4 million	18.0 million	20.0 million
\$22,000	13.6 million	17.0 million	19.0 million
\$24,000	13.2 million	16.5 million	18.5 million
\$26,000	12.8 million	16.0 million	18.0 million

Price and Demand Shifts: A Car Example

Now, imagine that the economy slows down so that many people lose their jobs or work fewer hours, reducing their incomes. In this case, the decrease in income would lead to a lower quantity of cars demanded at every given price, and the original demand curve D_0 would shift left to D_2 . The shift from D_0 to D_2 represents such a decrease in demand: At any given price level, the quantity demanded is now lower. In this example, a price of \$20,000 means 18 million cars sold along the original demand curve, but only 14.4 million sold after demand fell.

When a demand curve shifts, it does not mean that the quantity demanded by every individual buyer changes by the same amount. In this example, not everyone would have higher or lower income and not everyone would buy or not buy an additional car. Instead, a shift in a demand curve captures an pattern for the market as a whole.

In the previous section, we argued that higher income causes greater demand at every price. This is true for most goods and services. For some—

luxury cars, vacations in Europe, and fine jewelry—the effect of a rise in income can be especially pronounced. A product whose demand rises when income rises, and vice versa, is called a **normal good**. A few exceptions to this pattern do exist. As incomes rise, many people will buy fewer generic brand groceries and more name brand groceries. They are less likely to buy used cars and more likely to buy new cars. They will be less likely to rent an apartment and more likely to own a home, and so on. A product whose demand falls when income rises, and vice versa, is called an **inferior good**. In other words, when income increases, the demand curve shifts to the left.

Other Factors That Shift Demand Curves

Income is not the only factor that causes a shift in demand. Other things that change demand include tastes and preferences, the composition or size of the population, the prices of related goods, and even expectations. A change in any one of the underlying factors that determine what quantity people are willing to buy at a given price will cause a shift in demand. Graphically, the new demand curve lies either to the right (an increase) or to the left (a decrease) of the original demand curve. Let's look at these factors.

Changing Tastes or Preferences

From 1980 to 2014, the per-person consumption of chicken by Americans rose from 48 pounds per year to 85 pounds per year, and consumption of beef fell from 77 pounds per year to 54 pounds per year, according to the U.S. Department of Agriculture (USDA). Changes like these are largely due to movements in taste, which change the quantity of a good demanded at every price: that is, they shift the demand curve for that good, rightward for chicken and leftward for beef.

Changes in the Composition of the Population

The proportion of elderly citizens in the United States population is rising. It rose from 9.8% in 1970 to 12.6% in 2000, and will be projected (by the U.S. Census Bureau) 20% of the population by 2030. A society with relatively more children, like the United States in the 1960s, will have greater demand for goods and services like tricycles and day care facilities.

A society with relatively more elderly persons, as the United States is projected to have by 2030, has a higher demand for nursing homes and hearing aids. Similarly, changes in the size of the population can affect the demand for housing and many other goods. Each of these changes in demand will be shown as a shift in the demand curve.

The demand for a product can also be affected by changes in the prices of related goods such as substitutes or complements. A **substitute** is a good or service that can be used in place of another good or service. As electronic books, like this one, become more available, you would expect to see a decrease in demand for traditional printed books. A lower price for a substitute decreases demand for the other product. For example, in recent years as the price of tablet computers has fallen, the quantity demanded has increased (because of the law of demand). Since people are purchasing tablets, there has been a decrease in demand for laptops, which can be shown graphically as a leftward shift in the demand curve for laptops. A higher price for a substitute good has the reverse effect.

Other goods are **complements** for each other, meaning that the goods are often used together, because consumption of one good tends to enhance consumption of the other. Examples include breakfast cereal and milk; notebooks and pens or pencils, golf balls and golf clubs; gasoline and sport utility vehicles; and the five-way combination of bacon, lettuce, tomato, mayonnaise, and bread. If the price of golf clubs rises, since the quantity demanded of golf clubs falls (because of the law of demand), demand for a complement good like golf balls decreases, too. Similarly, a higher price for skis would shift the demand curve for a complement good like ski resort trips to the left, while a lower price for a complement has the reverse effect.

Changes in Expectations about Future Prices or Other Factors that Affect Demand

While it is clear that the price of a good affects the quantity demanded, it is also true that expectations about the future price (or expectations about tastes and preferences, income, and so on) can affect demand. For example, if people hear that a hurricane is coming, they may rush to the store to buy flashlight batteries and bottled water. If people learn that the price of a good like coffee is likely to rise in the future, they may head for the store to stock

up on coffee now. These changes in demand are shown as shifts in the curve. Therefore, a **shift in demand** happens when a change in some economic factor (other than price) causes a different quantity to be demanded at every price. The following Work It Out feature shows how this happens.

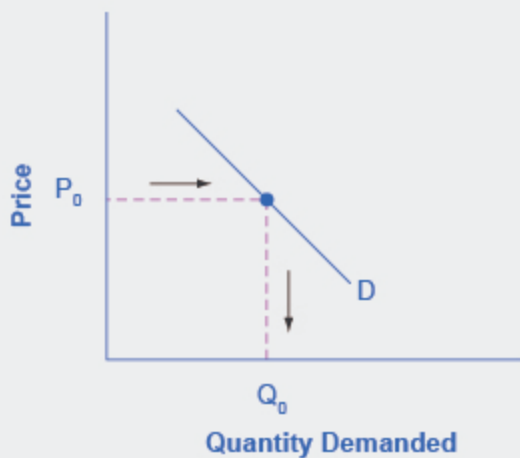
Note:

Shift in Demand

A shift in demand means that at any price (and at every price), the quantity demanded will be different than it was before. Following is an example of a shift in demand due to an income increase.

Step 1. Draw the graph of a demand curve for a normal good like pizza. Pick a price (like P_0). Identify the corresponding Q_0 . An example is shown in [\[link\]](#).

Demand Curve

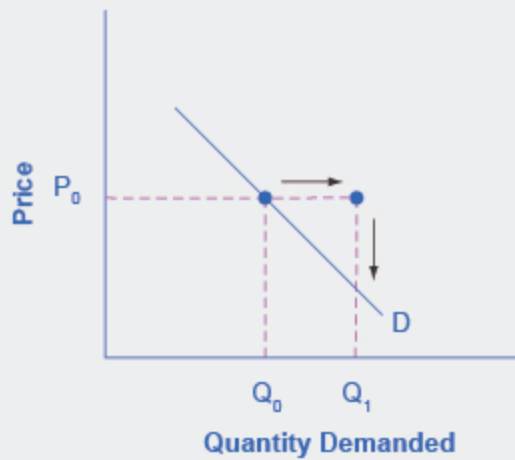


The demand curve can be used to identify how much consumers would buy at any given price.

Step 2. Suppose income increases. As a result of the change, are consumers going to buy more or less pizza? The answer is more. Draw a dotted horizontal line from the chosen price, through the original quantity demanded, to the new point with the new Q_1 . Draw a dotted vertical line

down to the horizontal axis and label the new Q_1 . An example is provided in [\[link\]](#).

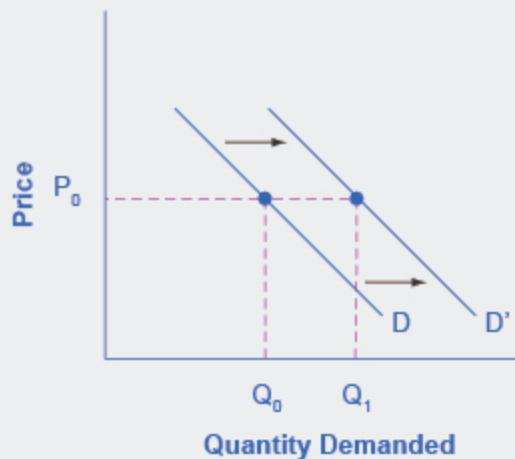
Demand Curve with Income Increase



With an increase in income, consumers will purchase larger quantities, pushing demand to the right.

Step 3. Now, shift the curve through the new point. You will see that an increase in income causes an upward (or rightward) shift in the demand curve, so that at any price the quantities demanded will be higher, as shown in [\[link\]](#).

Demand Curve Shifted Right



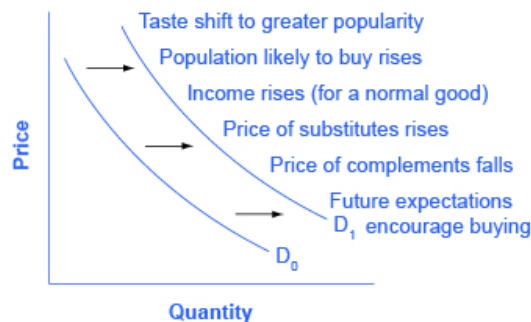
With an increase in income, consumers will

purchase larger quantities, pushing demand to the right, and causing the demand curve to shift right.

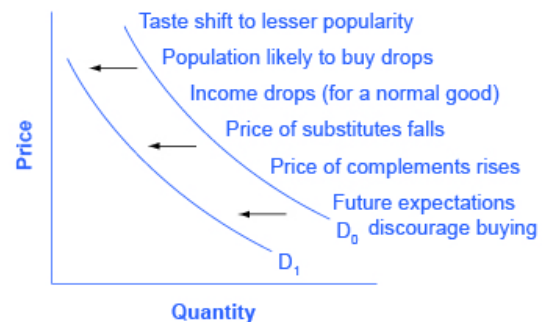
Summing Up Factors That Change Demand

Six factors that can shift demand curves are summarized in [\[link\]](#). The direction of the arrows indicates whether the demand curve shifts represent an increase in demand or a decrease in demand. Notice that a change in the price of the good or service itself is not listed among the factors that can shift a demand curve. A change in the price of a good or service causes a movement along a specific demand curve, and it typically leads to some change in the quantity demanded, but it does not shift the demand curve.

Factors That Shift Demand Curves



(a) Factors that increase demand



(b) Factors that decrease demand

- (a) A list of factors that can cause an increase in demand from D_0 to D_1 . (b) The same factors, if their direction is reversed, can cause a decrease in demand from D_0 to D_1 .

When a demand curve shifts, it will then intersect with a given supply curve at a different equilibrium price and quantity. We are, however, getting ahead of our story. Before discussing how changes in demand can affect

equilibrium price and quantity, we first need to discuss shifts in supply curves.

How Production Costs Affect Supply

A supply curve shows how quantity supplied will change as the price rises and falls, assuming *ceteris paribus* so that no other economically relevant factors are changing. If other factors relevant to supply do change, then the entire supply curve will shift. Just as a shift in demand is represented by a change in the quantity demanded at every price, a **shift in supply** means a change in the quantity supplied at every price.

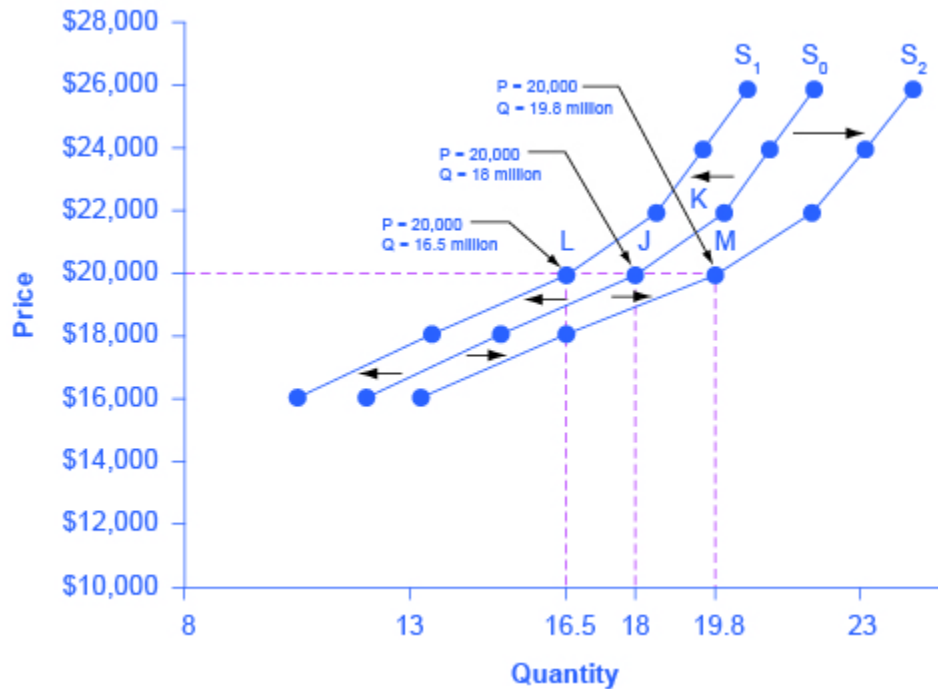
In thinking about the factors that affect supply, remember what motivates firms: profits, which are the difference between revenues and costs. Goods and services are produced using combinations of labor, materials, and machinery, or what we call **inputs** or **factors of production**. If a firm faces lower costs of production, while the prices for the good or service the firm produces remain unchanged, a firm's profits go up. When a firm's profits increase, it is more motivated to produce output, since the more it produces the more profit it will earn. So, when costs of production fall, a firm will tend to supply a larger quantity at any given price for its output. This can be shown by the supply curve shifting to the right.

Take, for example, a messenger company that delivers packages around a city. The company may find that buying gasoline is one of its main costs. If the price of gasoline falls, then the company will find it can deliver messages more cheaply than before. Since lower costs correspond to higher profits, the messenger company may now supply more of its services at any given price. For example, given the lower gasoline prices, the company can now serve a greater area, and increase its supply.

Conversely, if a firm faces higher costs of production, then it will earn lower profits at any given selling price for its products. As a result, a higher cost of production typically causes a firm to supply a smaller quantity at any given price. In this case, the supply curve shifts to the left.

Consider the supply for cars, shown by curve S_0 in [\[link\]](#). Point J indicates that if the price is \$20,000, the quantity supplied will be 18 million cars. If the price rises to \$22,000 per car, *ceteris paribus*, the quantity supplied will rise to 20 million cars, as point K on the S_0 curve shows. The same information can be shown in table form, as in [\[link\]](#).

Shifts in Supply: A Car Example



Decreased supply means that at every given price, the quantity supplied is lower, so that the supply curve shifts to the left, from S_0 to S_1 . Increased supply means that at every given price, the quantity supplied is higher, so that the supply curve shifts to the right, from S_0 to S_2 .

Price	Decrease to S_1	Original Quantity Supplied S_0	Increase to S_2
\$16,000	10.5 million	12.0 million	13.2 million
\$18,000	13.5 million	15.0 million	16.5 million
\$20,000	16.5 million	18.0 million	19.8 million
\$22,000	18.5 million	20.0 million	22.0 million
\$24,000	19.5 million	21.0 million	23.1 million
\$26,000	20.5 million	22.0 million	24.2 million

Price and Shifts in Supply: A Car Example

Now, imagine that the price of steel, an important ingredient in manufacturing cars, rises, so that producing a car has become more expensive. At any given price for selling cars, car manufacturers will react by supplying a lower quantity. This can be shown graphically as a leftward shift of supply, from S_0 to S_1 , which indicates that at any given price, the quantity supplied decreases. In this example, at a price of \$20,000, the quantity supplied decreases from 18 million on the original supply curve (S_0) to 16.5 million on the supply curve S_1 , which is labeled as point L.

Conversely, if the price of steel decreases, producing a car becomes less expensive. At any given price for selling cars, car manufacturers can now expect to earn higher profits, so they will supply a higher quantity. The shift of supply to the right, from S_0 to S_2 , means that at all prices, the quantity supplied has increased. In this example, at a price of \$20,000, the quantity

supplied increases from 18 million on the original supply curve (S_0) to 19.8 million on the supply curve S_2 , which is labeled M.

Other Factors That Affect Supply

In the example above, we saw that changes in the prices of inputs in the production process will affect the cost of production and thus the supply. Several other things affect the cost of production, too, such as changes in weather or other natural conditions, new technologies for production, and some government policies.

The cost of production for many agricultural products will be affected by changes in natural conditions. For example, in 2014 the Manchurian Plain in Northeastern China, which produces most of the country's wheat, corn, and soybeans, experienced its most severe drought in 50 years. A drought decreases the supply of agricultural products, which means that at any given price, a lower quantity will be supplied; conversely, especially good weather would shift the supply curve to the right.

When a firm discovers a new technology that allows the firm to produce at a lower cost, the supply curve will shift to the right, as well. For instance, in the 1960s a major scientific effort nicknamed the Green Revolution focused on breeding improved seeds for basic crops like wheat and rice. By the early 1990s, more than two-thirds of the wheat and rice in low-income countries around the world was grown with these Green Revolution seeds—and the harvest was twice as high per acre. A technological improvement that reduces costs of production will shift supply to the right, so that a greater quantity will be produced at any given price.

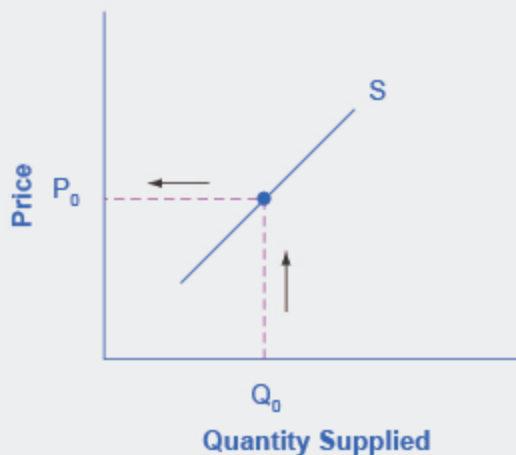
Government policies can affect the cost of production and the supply curve through taxes, regulations, and subsidies. For example, the U.S. government imposes a tax on alcoholic beverages that collects about \$8 billion per year from producers. Taxes are treated as costs by businesses. Higher costs decrease supply for the reasons discussed above. Other examples of policy that can affect cost are the wide array of government regulations that require firms to spend money to provide a cleaner environment or a safer workplace; complying with regulations increases costs.

A government subsidy, on the other hand, is the opposite of a tax. A subsidy occurs when the government pays a firm directly or reduces the firm's taxes if the firm carries out certain actions. From the firm's perspective, taxes or regulations are an additional cost of production that shifts supply to the left, leading the firm to produce a lower quantity at every given price. Government subsidies reduce the cost of production and increase supply at every given price, shifting supply to the right. The following Work It Out feature shows how this shift happens.

Note:**Shift in Supply**

We know that a supply curve shows the minimum price a firm will accept to produce a given quantity of output. What happens to the supply curve when the cost of production goes up? Following is an example of a shift in supply due to a production cost increase.

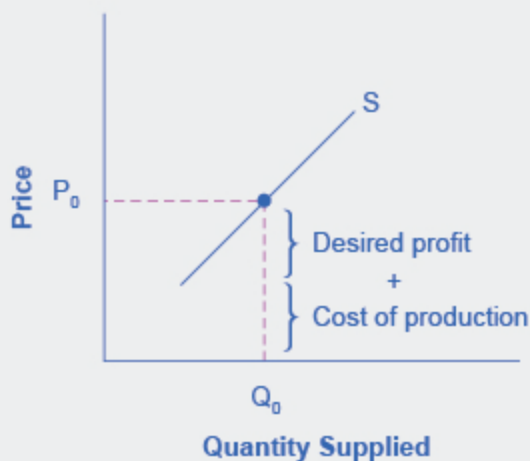
Step 1. Draw a graph of a supply curve for pizza. Pick a quantity (like Q_0). If you draw a vertical line up from Q_0 to the supply curve, you will see the price the firm chooses. An example is shown in [\[link\]](#).

Supply Curve

The supply curve can be used to show the minimum price a firm will accept to produce a given quantity of output.

Step 2. Why did the firm choose that price and not some other? One way to think about this is that the price is composed of two parts. The first part is the cost of producing pizzas at the margin; in this case, the cost of producing the pizza, including cost of ingredients (dough, sauce, cheese, pepperoni, and so on), the cost of the pizza oven, the rent on the shop, and the wages of the workers. The second part is the firm's desired profit, which is determined, among other factors, by the profit margins in that particular business. If you add these two parts together, you get the price the firm wishes to charge. The quantity Q_0 and associated price P_0 give you one point on the firm's supply curve, as shown in [\[link\]](#).

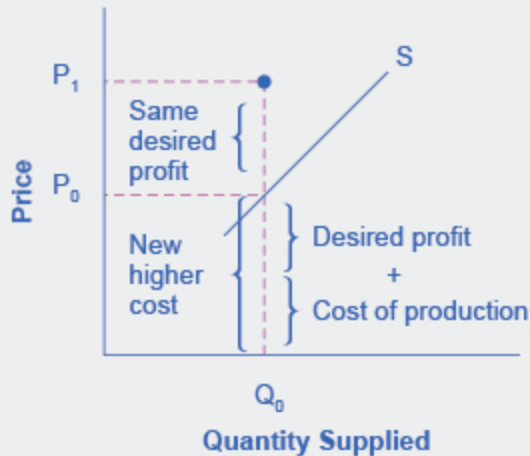
Setting Prices



The cost of production and the desired profit equal the price a firm will set for a product.

Step 3. Now, suppose that the cost of production goes up. Perhaps cheese has become more expensive by \$0.75 per pizza. If that is true, the firm will want to raise its price by the amount of the increase in cost (\$0.75). Draw this point on the supply curve directly above the initial point on the curve, but \$0.75 higher, as shown in [\[link\]](#).

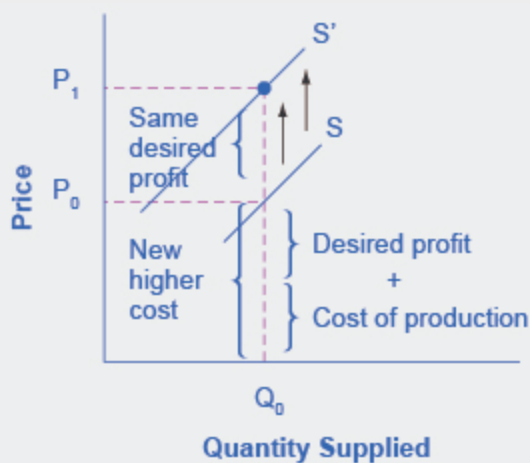
Increasing Costs Leads to Increasing Price



Because the cost of production and the desired profit equal the price a firm will set for a product, if the cost of production increases, the price for the product will also need to increase.

Step 4. Shift the supply curve through this point. You will see that an increase in cost causes an upward (or a leftward) shift of the supply curve so that at any price, the quantities supplied will be smaller, as shown in [\[link\]](#).

Supply Curve Shifts



When the cost of production increases, the supply curve shifts upwardly to a new price

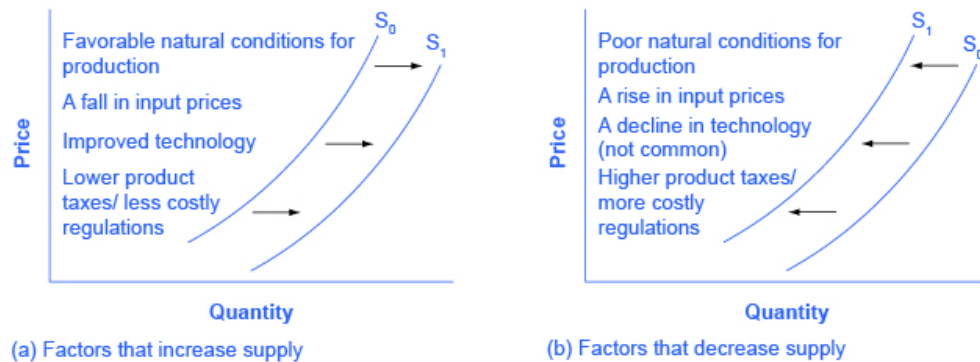
level.

Summing Up Factors That Change Supply

Changes in the cost of inputs, natural disasters, new technologies, and the impact of government decisions all affect the cost of production. In turn, these factors affect how much firms are willing to supply at any given price.

[\[link\]](#) summarizes factors that change the supply of goods and services. Notice that a change in the price of the product itself is not among the factors that shift the supply curve. Although a change in price of a good or service typically causes a change in quantity supplied or a movement along the supply curve for that specific good or service, it does not cause the supply curve itself to shift.

Factors That Shift Supply Curves



- (a) A list of factors that can cause an increase in supply from S_0 to S_1 .
- (b) The same factors, if their direction is reversed, can cause a decrease in supply from S_0 to S_1 .

Because demand and supply curves appear on a two-dimensional diagram with only price and quantity on the axes, an unwary visitor to the land of economics might be fooled into believing that economics is about only four

topics: demand, supply, price, and quantity. However, demand and supply are really “umbrella” concepts: demand covers all the factors that affect demand, and supply covers all the factors that affect supply. Factors other than price that affect demand and supply are included by using shifts in the demand or the supply curve. In this way, the two-dimensional demand and supply model becomes a powerful tool for analyzing a wide range of economic circumstances.

Key Concepts and Summary

Economists often use the *ceteris paribus* or “other things being equal” assumption: while examining the economic impact of one event, all other factors remain unchanged for the purpose of the analysis. Factors that can shift the demand curve for goods and services, causing a different quantity to be demanded at any given price, include changes in tastes, population, income, prices of substitute or complement goods, and expectations about future conditions and prices. Factors that can shift the supply curve for goods and services, causing a different quantity to be supplied at any given price, include input prices, natural conditions, changes in technology, and government taxes, regulations, or subsidies.

Self-Check Questions

Exercise:

Problem: Why do economists use the *ceteris paribus* assumption?

Solution:

To make it easier to analyze complex problems. *Ceteris paribus* allows you to look at the effect of one factor at a time on what it is you are trying to analyze. When you have analyzed all the factors individually, you add the results together to get the final answer.

Exercise:

Problem:

In an analysis of the market for paint, an economist discovers the facts listed below. State whether each of these changes will affect supply or demand, and in what direction.

- a. There have recently been some important cost-saving inventions in the technology for making paint.
 - b. Paint is lasting longer, so that property owners need not repaint as often.
 - c. Because of severe hailstorms, many people need to repaint now.
 - d. The hailstorms damaged several factories that make paint, forcing them to close down for several months.
-

Solution:

- a. An improvement in technology that reduces the cost of production will cause an increase in supply. Alternatively, you can think of this as a reduction in price necessary for firms to supply any quantity. Either way, this can be shown as a rightward (or downward) shift in the supply curve.
- b. An improvement in product quality is treated as an increase in tastes or preferences, meaning consumers demand more paint at any price level, so demand increases or shifts to the right. If this seems counterintuitive, note that demand in the future for the longer-lasting paint will fall, since consumers are essentially shifting demand from the future to the present.
- c. An increase in need causes an increase in demand or a rightward shift in the demand curve.
- d. Factory damage means that firms are unable to supply as much in the present. Technically, this is an increase in the cost of production. Either way you look at it, the supply curve shifts to the left.

Exercise:

Problem:

Many changes are affecting the market for oil. Predict how each of the following events will affect the equilibrium price and quantity in the market for oil. In each case, state how the event will affect the supply and demand diagram. Create a sketch of the diagram if necessary.

- a. Cars are becoming more fuel efficient, and therefore get more miles to the gallon.
- b. The winter is exceptionally cold.
- c. A major discovery of new oil is made off the coast of Norway.
- d. The economies of some major oil-using nations, like Japan, slow down.
- e. A war in the Middle East disrupts oil-pumping schedules.
- f. Landlords install additional insulation in buildings.
- g. The price of solar energy falls dramatically.
- h. Chemical companies invent a new, popular kind of plastic made from oil.

Solution:

- a. More fuel-efficient cars means there is less need for gasoline. This causes a leftward shift in the demand for gasoline and thus oil. Since the demand curve is shifting down the supply curve, the equilibrium price and quantity both fall.
- b. Cold weather increases the need for heating oil. This causes a rightward shift in the demand for heating oil and thus oil. Since the demand curve is shifting up the supply curve, the equilibrium price and quantity both rise.
- c. A discovery of new oil will make oil more abundant. This can be shown as a rightward shift in the supply curve, which will cause a decrease in the equilibrium price along with an increase in the equilibrium quantity. (The supply curve shifts down the demand curve so price and quantity follow the law of demand. If price goes down, then the quantity goes up.)

- d. When an economy slows down, it produces less output and demands less input, including energy, which is used in the production of virtually everything. A decrease in demand for energy will be reflected as a decrease in the demand for oil, or a leftward shift in demand for oil. Since the demand curve is shifting down the supply curve, both the equilibrium price and quantity of oil will fall.
- e. Disruption of oil pumping will reduce the supply of oil. This leftward shift in the supply curve will show a movement up the demand curve, resulting in an increase in the equilibrium price of oil and a decrease in the equilibrium quantity.
- f. Increased insulation will decrease the demand for heating. This leftward shift in the demand for oil causes a movement down the supply curve, resulting in a decrease in the equilibrium price and quantity of oil.
- g. Solar energy is a substitute for oil-based energy. So if solar energy becomes cheaper, the demand for oil will decrease as consumers switch from oil to solar. The decrease in demand for oil will be shown as a leftward shift in the demand curve. As the demand curve shifts down the supply curve, both equilibrium price and quantity for oil will fall.
- h. A new, popular kind of plastic will increase the demand for oil. The increase in demand will be shown as a rightward shift in demand, raising the equilibrium price and quantity of oil.

Review Questions

Exercise:

Problem:

When analyzing a market, how do economists deal with the problem that many factors that affect the market are changing at the same time?

Exercise:

Problem:

Name some factors that can cause a shift in the demand curve in markets for goods and services.

Exercise:**Problem:**

Name some factors that can cause a shift in the supply curve in markets for goods and services.

Critical Thinking Questions**Exercise:****Problem:**

Consider the demand for hamburgers. If the price of a substitute good (for example, hot dogs) increases and the price of a complement good (for example, hamburger buns) increases, can you tell for sure what will happen to the demand for hamburgers? Why or why not? Illustrate your answer with a graph.

Exercise:**Problem:**

How do you suppose the demographics of an aging population of “Baby Boomers” in the United States will affect the demand for milk? Justify your answer.

Exercise:**Problem:**

We know that a change in the price of a product causes a movement along the demand curve. Suppose consumers believe that prices will be rising in the future. How will that affect demand for the product in the present? Can you show this graphically?

Exercise:**Problem:**

Suppose there is soda tax to curb obesity. What should a reduction in the soda tax do to the supply of sodas and to the equilibrium price and quantity? Can you show this graphically? *Hint:* assume that the soda tax is collected from the sellers

Problems**Exercise:****Problem:**

[\[link\]](#) shows information on the demand and supply for bicycles, where the quantities of bicycles are measured in thousands.

Price	Qd	Qs
\$120	50	36
\$150	40	40
\$180	32	48
\$210	28	56
\$240	24	70

- a. What is the quantity demanded and the quantity supplied at a price of \$210?

- b. At what price is the quantity supplied equal to 48,000?
- c. Graph the demand and supply curve for bicycles. How can you determine the equilibrium price and quantity from the graph? How can you determine the equilibrium price and quantity from the table? What are the equilibrium price and equilibrium quantity?
- d. If the price was \$120, what would the quantities demanded and supplied be? Would a shortage or surplus exist? If so, how large would the shortage or surplus be?

Exercise:

Problem:

The computer market in recent years has seen many more computers sell at much lower prices. What shift in demand or supply is most likely to explain this outcome? Sketch a demand and supply diagram and explain your reasoning for each.

- a. A rise in demand
- b. A fall in demand
- c. A rise in supply
- d. A fall in supply

References

Landsburg, Steven E. *The Armchair Economist: Economics and Everyday Life*. New York: The Free Press. 2012. specifically Section IV: How Markets Work.

National Chicken Council. 2015. "Per Capita Consumption of Poultry and Livestock, 1965 to Estimated 2015, in Pounds." Accessed April 13, 2015. <http://www.nationalchickencouncil.org/about-the-industry/statistics/per-capita-consumption-of-poultry-and-livestock-1965-to-estimated-2012-in-pounds/>.

Wessel, David. "Saudi Arabia Fears \$40-a-Barrel Oil, Too." *The Wall Street Journal*. May 27, 2004, p. 42.
<http://online.wsj.com/news/articles/SB108561000087822300>.

Glossary

ceteris paribus

other things being equal

complements

goods that are often used together so that consumption of one good tends to enhance consumption of the other

factors of production

the combination of labor, materials, and machinery that is used to produce goods and services; also called inputs

inferior good

a good in which the quantity demanded falls as income rises, and in which quantity demanded rises as income falls

inputs

the combination of labor, materials, and machinery that is used to produce goods and services; also called factors of production

normal good

a good in which the quantity demanded rises as income rises, and in which quantity demanded falls as income falls

shift in demand

when a change in some economic factor (other than price) causes a different quantity to be demanded at every price

shift in supply

when a change in some economic factor (other than price) causes a different quantity to be supplied at every price

substitute

a good that can replace another to some extent, so that greater consumption of one good can mean less of the other

Changes in Equilibrium Price and Quantity: The Four-Step Process

By the end of this section, you will be able to:

- Identify equilibrium price and quantity through the four-step process
- Graph equilibrium price and quantity
- Contrast shifts of demand or supply and movements along a demand or supply curve
- Graph demand and supply curves, including equilibrium price and quantity, based on real-world examples

Let's begin this discussion with a single economic event. It might be an event that affects demand, like a change in income, population, tastes, prices of substitutes or complements, or expectations about future prices. It might be an event that affects supply, like a change in natural conditions, input prices, or technology, or government policies that affect production. How does this economic event affect equilibrium price and quantity? We will analyze this question using a four-step process.

Step 1. Draw a demand and supply model before the economic change took place. To establish the model requires four standard pieces of information: The law of demand, which tells us the slope of the demand curve; the law of supply, which gives us the slope of the supply curve; the shift variables for demand; and the shift variables for supply. From this model, find the initial equilibrium values for price and quantity.

Step 2. Decide whether the economic change being analyzed affects demand or supply. In other words, does the event refer to something in the list of demand factors or supply factors?

Step 3. Decide whether the effect on demand or supply causes the curve to shift to the right or to the left, and sketch the new demand or supply curve on the diagram. In other words, does the event increase or decrease the amount consumers want to buy or producers want to sell?

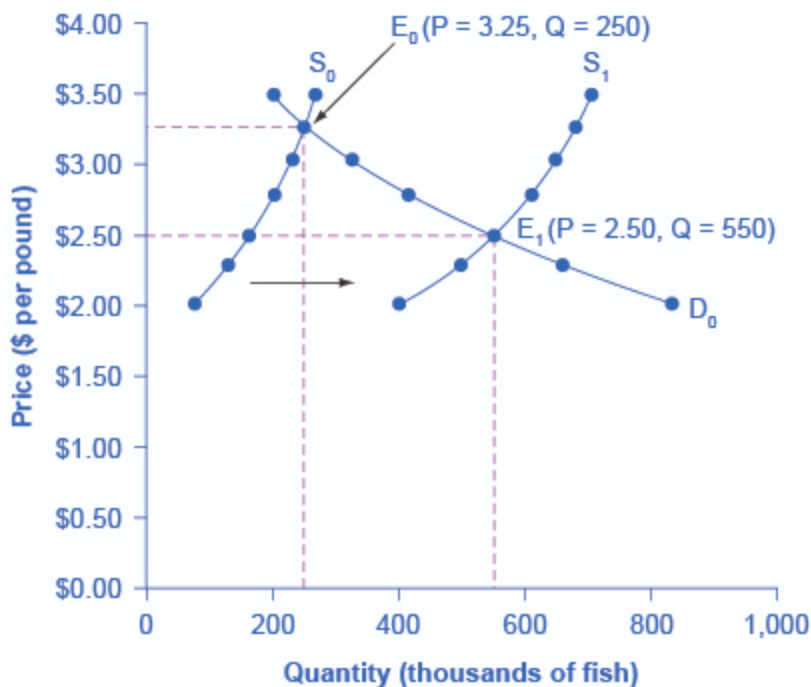
Step 4. Identify the new equilibrium and then compare the original equilibrium price and quantity to the new equilibrium price and quantity.

Let's consider one example that involves a shift in supply and one that involves a shift in demand. Then we will consider an example where both supply and demand shift.

Good Weather for Salmon Fishing

In the summer of 2000, weather conditions were excellent for commercial salmon fishing off the California coast. Heavy rains meant higher than normal levels of water in the rivers, which helps the salmon to breed. Slightly cooler ocean temperatures stimulated the growth of plankton, the microscopic organisms at the bottom of the ocean food chain, providing everything in the ocean with a hearty food supply. The ocean stayed calm during fishing season, so commercial fishing operations did not lose many days to bad weather. How did these climate conditions affect the quantity and price of salmon? [\[link\]](#) illustrates the four-step approach, which is explained below, to work through this problem. [\[link\]](#) provides the information to work the problem as well.

Good Weather for Salmon Fishing: The Four-Step Process



Unusually good weather leads to changes in the price and quantity of salmon.

Price per Pound	Quantity Supplied in 1999	Quantity Supplied in 2000	Quantity Demanded
\$2.00	80	400	840
\$2.25	120	480	680
\$2.50	160	550	550
\$2.75	200	600	450
\$3.00	230	640	350
\$3.25	250	670	250
\$3.50	270	700	200

Salmon Fishing

Step 1. Draw a demand and supply model to illustrate the market for salmon in the year before the good weather conditions began. The demand curve D_0 and the supply curve S_0 show that the original equilibrium price is \$3.25 per pound and the original equilibrium quantity is 250,000 fish. (This price per pound is what commercial buyers pay at the fishing docks; what consumers pay at the grocery is higher.)

Step 2. Did the economic event affect supply or demand? Good weather is an example of a natural condition that affects supply.

Step 3. Was the effect on supply an increase or a decrease? Good weather is a change in natural conditions that increases the quantity supplied at any given price. The supply curve shifts to the right, moving from the original supply curve S_0 to the new supply curve S_1 , which is shown in both the table and the figure.

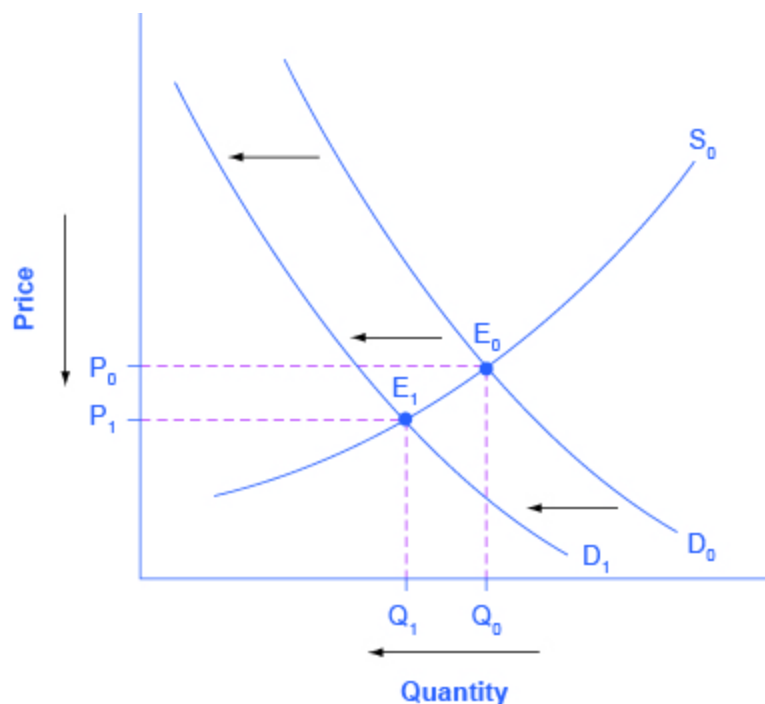
Step 4. Compare the new equilibrium price and quantity to the original equilibrium. At the new equilibrium E_1 , the equilibrium price falls from \$3.25 to \$2.50, but the equilibrium quantity increases from 250,000 to 550,000 salmon. Notice that the equilibrium quantity demanded increased, even though the demand curve did not move.

In short, good weather conditions increased supply of the California commercial salmon. The result was a higher equilibrium quantity of salmon bought and sold in the market at a lower price.

Newspapers and the Internet

According to the Pew Research Center for People and the Press, more and more people, especially younger people, are getting their news from online and digital sources. The majority of U.S. adults now own smartphones or tablets, and most of those Americans say they use them in part to get the news. From 2004 to 2012, the share of Americans who reported getting their news from digital sources increased from 24% to 39%. How has this affected consumption of print news media, and radio and television news? [\[link\]](#) and the text below illustrates using the four-step analysis to answer this question.

The Print News Market: A Four-Step Analysis



A change in tastes from print news sources to digital sources results in a leftward shift in demand for the former. The result is a decrease in both equilibrium price and quantity.

Step 1. Develop a demand and supply model to think about what the market looked like before the event. The demand curve D_0 and the supply curve S_0 show the original relationships. In this case, the analysis is performed without specific numbers on the price and quantity axis.

Step 2. Did the change described affect supply or demand? A change in tastes, from traditional news sources (print, radio, and television) to digital sources, caused a change in demand for the former.

Step 3. Was the effect on demand positive or negative? A shift to digital news sources will tend to mean a lower quantity demanded of traditional news sources at every given price, causing the demand curve for print and other traditional news sources to shift to the left, from D_0 to D_1 .

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The decline in print news reading predates 2004. Print newspaper circulation peaked in 1973 and has declined since then due to competition from television and radio news. In 1991, 55% of Americans indicated they got their news from print sources, while only 29% did so in 2012. Radio news has followed a similar path in recent decades, with the share of Americans getting their news from radio declining from 54% in 1991 to 33% in 2012. Television news has held its own over the last 15 years, with a market share staying in the mid to upper fifties. What does this suggest for the future, given that two-thirds of Americans under 30 years old say they do not get their news from television at all?

The Interconnections and Speed of Adjustment in Real Markets

In the real world, many factors that affect demand and supply can change all at once. For example, the demand for cars might increase because of rising incomes and population, and it might decrease because of rising gasoline prices (a complementary good). Likewise, the supply of cars might increase because of innovative new technologies that reduce the cost of car production, and it might decrease as a result of new government regulations requiring the installation of costly pollution-control technology.

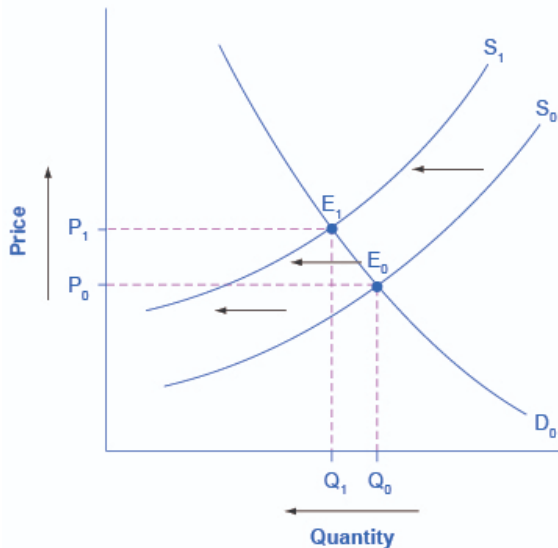
Moreover, rising incomes and population or changes in gasoline prices will affect many markets, not just cars. How can an economist sort out all these interconnected events? The answer lies in the *ceteris paribus* assumption. Look at how each economic event affects each market, one event at a time, holding all else constant. Then combine the analyses to see the net effect.

A Combined Example

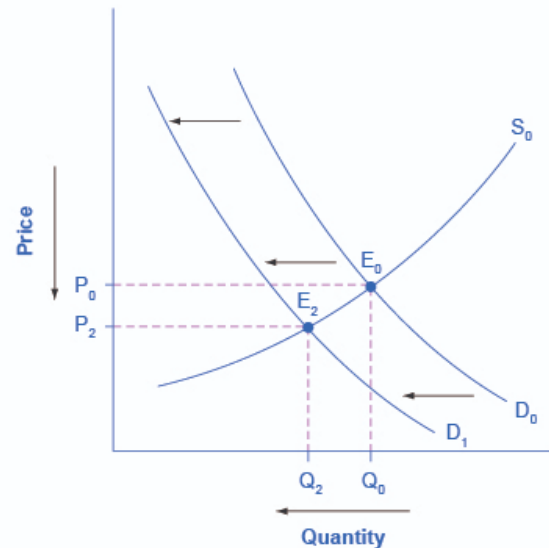
The U.S. Postal Service is facing difficult challenges. Compensation for postal workers tends to increase most years due to cost-of-living increases.

At the same time, more and more people are using email, text, and other digital message forms such as Facebook and Twitter to communicate with friends and others. What does this suggest about the continued viability of the Postal Service? [\[link\]](#) and the text below illustrates using the four-step analysis to answer this question.

Higher Compensation for Postal Workers: A Four-Step Analysis



(a) Shift in supply



(b) Shift in demand

(a) Higher labor compensation causes a leftward shift in the supply curve, a decrease in the equilibrium quantity, and an increase in the equilibrium price. (b) A change in tastes away from Postal Services causes a leftward shift in the demand curve, a decrease in the equilibrium quantity, and a decrease in the equilibrium price.

Since this problem involves two disturbances, we need two four-step analyses, the first to analyze the effects of higher compensation for postal workers, the second to analyze the effects of many people switching from “snailmail” to email and other digital messages.

[\[link\]](#) (a) shows the shift in supply discussed in the following steps.

Step 1. Draw a demand and supply model to illustrate what the market for the U.S. Postal Service looked like before this scenario starts. The demand

curve D_0 and the supply curve S_0 show the original relationships.

Step 2. Did the change described affect supply or demand? Labor compensation is a cost of production. A change in production costs caused a change in supply for the Postal Service.

Step 3. Was the effect on supply positive or negative? Higher labor compensation leads to a lower quantity supplied of postal services at every given price, causing the supply curve for postal services to shift to the left, from S_0 to S_1 .

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a higher price than the original equilibrium (E_0).

[\[link\]](#) (b) shows the shift in demand discussed in the following steps.

Step 1. Draw a demand and supply model to illustrate what the market for U.S. Postal Services looked like before this scenario starts. The demand curve D_0 and the supply curve S_0 show the original relationships. Note that this diagram is independent from the diagram in panel (a).

Step 2. Did the change described affect supply or demand? A change in tastes away from snailmail toward digital messages will cause a change in demand for the Postal Service.

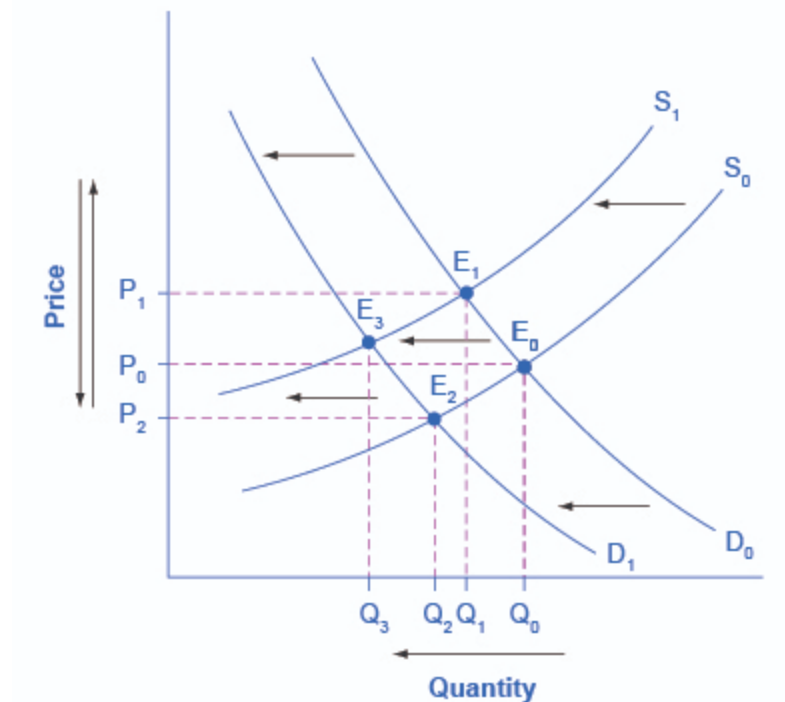
Step 3. Was the effect on demand positive or negative? A change in tastes away from snailmail toward digital messages causes lower quantity demanded of postal services at every given price, causing the demand curve for postal services to shift to the left, from D_0 to D_1 .

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_2) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The final step in a scenario where both supply and demand shift is to combine the two individual analyses to determine what happens to the

equilibrium quantity and price. Graphically, we superimpose the previous two diagrams one on top of the other, as in [\[link\]](#).

Combined Effect of Decreased Demand and Decreased Supply



Supply and demand shifts cause changes in equilibrium price and quantity.

Following are the results:

Effect on Quantity: The effect of higher labor compensation on Postal Services because it raises the cost of production is to decrease the equilibrium quantity. The effect of a change in tastes away from snailmail is to decrease the equilibrium quantity. Since both shifts are to the left, the overall impact is a decrease in the equilibrium quantity of Postal Services (Q_3). This is easy to see graphically, since Q_3 is to the left of Q_0 .

Effect on Price: The overall effect on price is more complicated. The effect of higher labor compensation on Postal Services, because it raises the cost of production, is to increase the equilibrium price. The effect of a change in

tastes away from snailmail is to decrease the equilibrium price. Since the two effects are in opposite directions, unless we know the magnitudes of the two effects, the overall effect is unclear. This is not unusual. When both curves shift, typically we can determine the overall effect on price or on quantity, but not on both. In this case, we determined the overall effect on the equilibrium quantity, but not on the equilibrium price. In other cases, it might be the opposite.

The next Clear It Up feature focuses on the difference between shifts of supply or demand and movements along a curve.

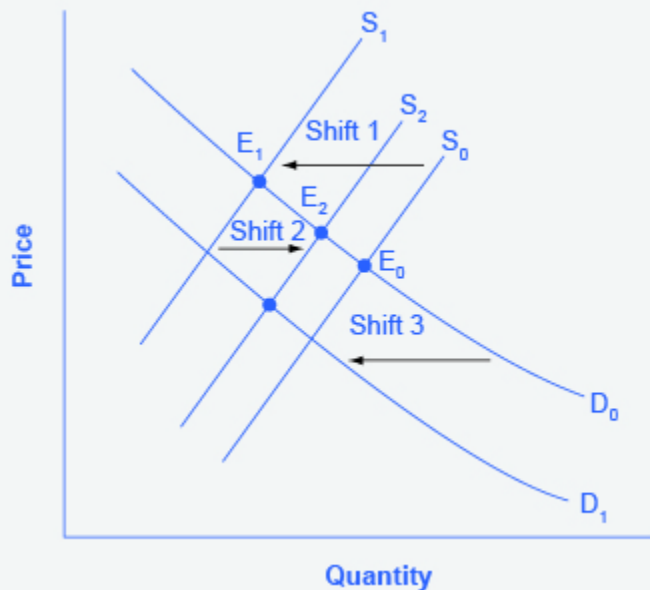
Note:

What is the difference between shifts of demand or supply versus movements along a demand or supply curve?

One common mistake in applying the demand and supply framework is to confuse the shift of a demand or a supply curve with movement along a demand or supply curve. As an example, consider a problem that asks whether a drought will increase or decrease the equilibrium quantity and equilibrium price of wheat. Lee, a student in an introductory economics class, might reason:

“Well, it is clear that a drought reduces supply, so I will shift back the supply curve, as in the shift from the original supply curve S_0 to S_1 shown on the diagram (called Shift 1). So the equilibrium moves from E_0 to E_1 , the equilibrium quantity is lower and the equilibrium price is higher. Then, a higher price makes farmers more likely to supply the good, so the supply curve shifts right, as shown by the shift from S_1 to S_2 , on the diagram (shown as Shift 2), so that the equilibrium now moves from E_1 to E_2 . The higher price, however, also reduces demand and so causes demand to shift back, like the shift from the original demand curve, D_0 to D_1 on the diagram (labeled Shift 3), and the equilibrium moves from E_2 to E_3 .”

Shifts of Demand or Supply versus Movements along a Demand or Supply Curve



A shift in one curve never causes a shift in the other curve. Rather, a shift in one curve causes a movement along the second curve.

At about this point, Lee suspects that this answer is headed down the wrong path. Think about what might be wrong with Lee's logic, and then read the answer that follows.

Answer: Lee's first step is correct: that is, a drought shifts back the supply curve of wheat and leads to a prediction of a lower equilibrium quantity and a higher equilibrium price. This corresponds to a movement along the original demand curve (D_0), from E_0 to E_1 . The rest of Lee's argument is wrong, because it mixes up shifts in supply with quantity supplied, and shifts in demand with quantity demanded. A higher or lower price never shifts the supply curve, as suggested by the shift in supply from S_1 to S_2 . Instead, a price change leads to a movement along a given supply curve. Similarly, a higher or lower price never shifts a demand curve, as suggested in the shift from D_0 to D_1 . Instead, a price change leads to a movement along a given demand curve. Remember, a change in the price of a good never causes the demand or supply curve for that good to shift. Think carefully about the timeline of events: What happens first, what happens next? What is cause, what is effect? If you keep the order right, you are more likely to get the analysis correct.

In the four-step analysis of how economic events affect equilibrium price and quantity, the movement from the old to the new equilibrium seems immediate. As a practical matter, however, prices and quantities often do not zoom straight to equilibrium. More realistically, when an economic event causes demand or supply to shift, prices and quantities set off in the general direction of equilibrium. Indeed, even as they are moving toward one new equilibrium, prices are often then pushed by another change in demand or supply toward another equilibrium.

Key Concepts and Summary

When using the supply and demand framework to think about how an event will affect the equilibrium price and quantity, proceed through four steps: (1) sketch a supply and demand diagram to think about what the market looked like before the event; (2) decide whether the event will affect supply or demand; (3) decide whether the effect on supply or demand is negative or positive, and draw the appropriate shifted supply or demand curve; (4) compare the new equilibrium price and quantity to the original ones.

Self-Check Questions

Exercise:

Problem:

Let's think about the market for air travel. From August 2014 to January 2015, the price of jet fuel decreased roughly 47%. Using the four-step analysis, how do you think this fuel price decrease affected the equilibrium price and quantity of air travel?

Solution:

Step 1. Draw the graph with the initial supply and demand curves. Label the initial equilibrium price and quantity.

Step 2. Did the economic event affect supply or demand? Jet fuel is a cost of producing air travel, so an increase in jet fuel price affects

supply.

Step 3. An increase in the price of jet fuel caused a decrease in the cost of air travel. We show this as a downward or rightward shift in supply.

Step 4. A rightward shift in supply causes a movement down the demand curve, lowering the equilibrium price of air travel and increasing the equilibrium quantity.

Exercise:

Problem:

A tariff is a tax on imported goods. Suppose the U.S. government cuts the tariff on imported flat screen televisions. Using the four-step analysis, how do you think the tariff reduction will affect the equilibrium price and quantity of flat screen TVs?

Solution:

Step 1. Draw the graph with the initial supply and demand curves. Label the initial equilibrium price and quantity.

Step 2. Did the economic event affect supply or demand? A tariff is treated like a cost of production, so this affects supply.

Step 3. A tariff reduction is equivalent to a decrease in the cost of production, which we can show as a rightward (or downward) shift in supply.

Step 4. A rightward shift in supply causes a movement down the demand curve, lowering the equilibrium price and raising the equilibrium quantity.

Review Questions

Exercise:

Problem:

How does one analyze a market where both demand and supply shift?

Exercise:**Problem:**

What causes a movement along the demand curve? What causes a movement along the supply curve?

Critical Thinking Questions**Exercise:****Problem:**

Use the four-step process to analyze the impact of the advent of the iPod (or other portable digital music players) on the equilibrium price and quantity of the Sony Walkman (or other portable audio cassette players).

Exercise:**Problem:**

Use the four-step process to analyze the impact of a reduction in tariffs on imports of iPods on the equilibrium price and quantity of Sony Walkman-type products.

Exercise:**Problem:**

Suppose both of these events took place at the same time. Combine your analyses of the impacts of the iPod and the tariff reduction to determine the likely impact on the equilibrium price and quantity of Sony Walkman-type products. Show your answer graphically.

Problems

Exercise:**Problem:**

Demand and supply in the market for cheddar cheese is illustrated in [\[link\]](#). Graph the data and find the equilibrium. Next, create a table showing the change in quantity demanded or quantity supplied, and a graph of the new equilibrium, in each of the following situations:

- a. The price of milk, a key input for cheese production, rises, so that the supply decreases by 80 pounds at every price.
- b. A new study says that eating cheese is good for your health, so that demand increases by 20% at every price.

Price per Pound	Qd	Qs
\$3.00	750	540
\$3.20	700	600
\$3.40	650	650
\$3.60	620	700
\$3.80	600	720
\$4.00	590	730

Exercise:

Problem:

Supply and demand for movie tickets in a city are shown in [\[link\]](#). Graph demand and supply and identify the equilibrium. Then calculate in a table and graph the effect of the following two changes.

- a. Three new nightclubs open. They offer decent bands and have no cover charge, but make their money by selling food and drink. As a result, demand for movie tickets falls by six units at every price.
- b. The city eliminates a tax that it had been placing on all local entertainment businesses. The result is that the quantity supplied of movies at any given price increases by 10%.

Price per Pound	Qd	Qs
\$5.00	26	16
\$6.00	24	18
\$7.00	22	20
\$8.00	21	21
\$9.00	20	22

References

Pew Research Center. "Pew Research: Center for the People & the Press."
<http://www.people-press.org/>.

Price Ceilings and Price Floors

By the end of this section, you will be able to:

- Explain price controls, price ceilings, and price floors
- Analyze demand and supply as a social adjustment mechanism

Controversy sometimes surrounds the prices and quantities established by demand and supply, especially for products that are considered necessities. In some cases, discontent over prices turns into public pressure on politicians, who may then pass legislation to prevent a certain price from climbing “too high” or falling “too low.”

The demand and supply model shows how people and firms will react to the incentives provided by these laws to control prices, in ways that will often lead to undesirable consequences. Alternative policy tools can often achieve the desired goals of price control laws, while avoiding at least some of their costs and tradeoffs.

Price Ceilings

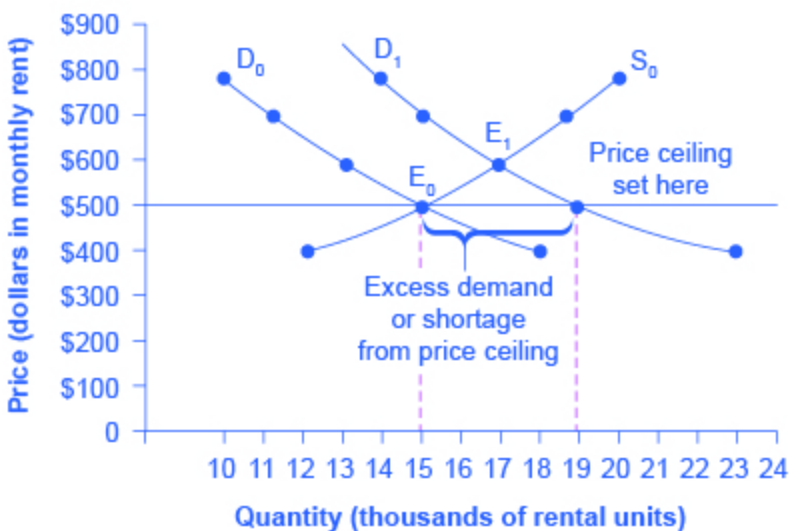
Laws that government enacts to regulate prices are called **Price controls**. Price controls come in two flavors. A **price ceiling** keeps a price from rising above a certain level (the “ceiling”), while a **price floor** keeps a price from falling below a certain level (the “floor”). This section uses the demand and supply framework to analyze price ceilings. The next section discusses price floors.

In many markets for goods and services, demanders outnumber suppliers. Consumers, who are also potential voters, sometimes unite behind a political proposal to hold down a certain price. In some cities, such as Albany, renters have pressed political leaders to pass rent control laws, a price ceiling that usually works by stating that rents can be raised by only a certain maximum percentage each year.

Rent control becomes a politically hot topic when rents begin to rise rapidly. Everyone needs an affordable place to live. Perhaps a change in tastes makes a certain suburb or town a more popular place to live. Perhaps

locally-based businesses expand, bringing higher incomes and more people into the area. Changes of this sort can cause a change in the demand for rental housing, as [\[link\]](#) illustrates. The original equilibrium (E_0) lies at the intersection of supply curve S_0 and demand curve D_0 , corresponding to an equilibrium price of \$500 and an equilibrium quantity of 15,000 units of rental housing. The effect of greater income or a change in tastes is to shift the demand curve for rental housing to the right, as shown by the data in [\[link\]](#) and the shift from D_0 to D_1 on the graph. In this market, at the new equilibrium E_1 , the price of a rental unit would rise to \$600 and the equilibrium quantity would increase to 17,000 units.

A Price Ceiling Example—Rent Control



The original intersection of demand and supply occurs at E_0 . If demand shifts from D_0 to D_1 , the new equilibrium would be at E_1 —unless a price ceiling prevents the price from rising. If the price is not permitted to rise, the quantity supplied remains at 15,000. However, after the change in demand, the quantity demanded rises to 19,000, resulting in a shortage.

Price	Original Quantity Supplied	Original Quantity Demanded	New Quantity Demanded
\$400	12,000	18,000	23,000
\$500	15,000	15,000	19,000
\$600	17,000	13,000	17,000
\$700	19,000	11,000	15,000
\$800	20,000	10,000	14,000

Rent Control

Suppose that a rent control law is passed to keep the price at the original equilibrium of \$500 for a typical apartment. In [\[link\]](#), the horizontal line at the price of \$500 shows the legally fixed maximum price set by the rent control law. However, the underlying forces that shifted the demand curve to the right are still there. At that price (\$500), the quantity supplied remains at the same 15,000 rental units, but the quantity demanded is 19,000 rental units. In other words, the quantity demanded exceeds the quantity supplied, so there is a shortage of rental housing. One of the ironies of price ceilings is that while the price ceiling was intended to help renters, there are actually fewer apartments rented out under the price ceiling (15,000 rental units) than would be the case at the market rent of \$600 (17,000 rental units).

Price ceilings do not simply benefit renters at the expense of landlords. Rather, some renters (or potential renters) lose their housing as landlords convert apartments to co-ops and condos. Even when the housing remains in the rental market, landlords tend to spend less on maintenance and on essentials like heating, cooling, hot water, and lighting. The first rule of economics is you do not get something for nothing—everything has an opportunity cost. So if renters get “cheaper” housing than the market requires, they tend to also end up with lower quality housing.

Price ceilings have been proposed for other products. For example, price ceilings to limit what producers can charge have been proposed in recent years for prescription drugs, doctor and hospital fees, the charges made by some automatic teller bank machines, and auto insurance rates. Price ceilings are enacted in an attempt to keep prices low for those who demand the product. But when the market price is not allowed to rise to the equilibrium level, quantity demanded exceeds quantity supplied, and thus a shortage occurs. Those who manage to purchase the product at the lower price given by the price ceiling will benefit, but sellers of the product will suffer, along with those who are not able to purchase the product at all. Quality is also likely to deteriorate.

Price Floors

A price floor is the lowest legal price that can be paid in markets for goods and services, labor, or financial capital. Perhaps the best-known example of a price floor is the minimum wage, which is based on the normative view that someone working full time ought to be able to afford a basic standard of living. The federal minimum wage at the end of 2014 was \$7.25 per hour, which yields an income for a single person slightly higher than the poverty line. As the cost of living rises over time, the Congress periodically raises the federal minimum wage.

Price floors are sometimes called “price supports,” because they support a price by preventing it from falling below a certain level. Around the world, many countries have passed laws to create agricultural price supports. Farm prices and thus farm incomes fluctuate, sometimes widely. So even if, on average, farm incomes are adequate, some years they can be quite low. The purpose of price supports is to prevent these swings.

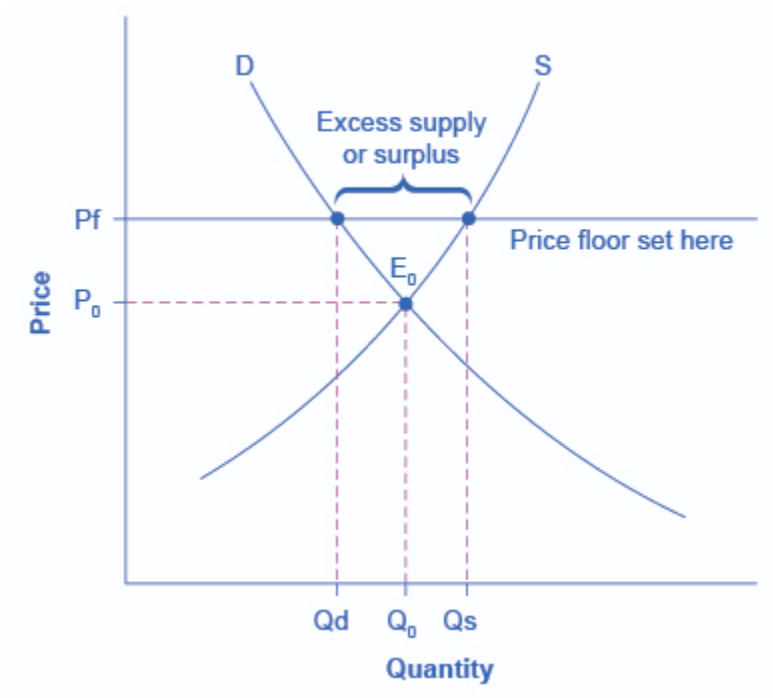
The most common way price supports work is that the government enters the market and buys up the product, adding to demand to keep prices higher than they otherwise would be. According to the Common Agricultural Policy reform passed in 2013, the European Union (EU) will spend about 60 billion euros per year, or 67 billion dollars per year, or roughly 38% of the EU budget, on price supports for Europe’s farmers from 2014 to 2020.

[\[link\]](#) illustrates the effects of a government program that assures a price above the equilibrium by focusing on the market for wheat in Europe. In the absence of government intervention, the price would adjust so that the quantity supplied would equal the quantity demanded at the equilibrium point E_0 , with price P_0 and quantity Q_0 . However, policies to keep prices high for farmers keeps the price above what would have been the market equilibrium level—the price P_f shown by the dashed horizontal line in the diagram. The result is a quantity supplied in excess of the quantity demanded (Q_d). When quantity supplied exceeds quantity demanded, a surplus exists.

The high-income areas of the world, including the United States, Europe, and Japan, are estimated to spend roughly \$1 billion per day in supporting their farmers. If the government is willing to purchase the excess supply (or to provide payments for others to purchase it), then farmers will benefit from the price floor, but taxpayers and consumers of food will pay the costs. Numerous proposals have been offered for reducing farm subsidies. In many countries, however, political support for subsidies for farmers remains strong. Either because this is viewed by the population as supporting the traditional rural way of life or because of the lobbying power of the agribusiness industry.

For more detail on the effects price ceilings and floors have on demand and supply, see the following Clear It Up feature.

European Wheat Prices: A Price Floor Example



The intersection of demand (D) and supply (S) would be at the equilibrium point E_0 .

However, a price floor set at P_f holds the price above E_0 and prevents it from falling.

The result of the price floor is that the quantity supplied Q_s exceeds the quantity demanded Q_d . There is excess supply, also called a surplus.

Note:

Do price ceilings and floors change demand or supply?

Neither price ceilings nor price floors cause demand or supply to change. They simply set a price that limits what can be legally charged in the market. Remember, changes in price do not cause demand or supply to change. Price ceilings and price floors can cause a different choice of quantity demanded along a demand curve, but they do not move the

demand curve. Price controls can cause a different choice of quantity supplied along a supply curve, but they do not shift the supply curve.

Key Concepts and Summary

Price ceilings prevent a price from rising above a certain level. When a price ceiling is set below the equilibrium price, quantity demanded will exceed quantity supplied, and excess demand or shortages will result. Price floors prevent a price from falling below a certain level. When a price floor is set above the equilibrium price, quantity supplied will exceed quantity demanded, and excess supply or surpluses will result. Price floors and price ceilings often lead to unintended consequences.

Self-Check Questions

Exercise:

Problem:

What is the effect of a price ceiling on the quantity demanded of the product? What is the effect of a price ceiling on the quantity supplied? Why exactly does a price ceiling cause a shortage?

Solution:

A price ceiling (which is below the equilibrium price) will cause the quantity demanded to rise and the quantity supplied to fall. This is why a price ceiling creates a shortage.

Exercise:

Problem: Does a price ceiling change the equilibrium price?

Solution:

A price ceiling is just a legal restriction. Equilibrium is an economic condition. People may or may not obey the price ceiling, so the actual price may be at or above the price ceiling, but the price ceiling does not change the equilibrium price.

Exercise:

Problem:

What would be the impact of imposing a price floor below the equilibrium price?

Solution:

A price ceiling is a legal maximum price, but a price floor is a legal minimum price and, consequently, it would leave room for the price to rise to its equilibrium level. In other words, a price floor below equilibrium will not be binding and will have no effect.

Review Questions

Exercise:

Problem:

Does a price ceiling attempt to make a price higher or lower?

Exercise:

Problem:

How does a price ceiling set below the equilibrium level affect quantity demanded and quantity supplied?

Exercise:

Problem: Does a price floor attempt to make a price higher or lower?

Exercise:

Problem:

How does a price floor set above the equilibrium level affect quantity demanded and quantity supplied?

Critical Thinking Questions**Exercise:****Problem:**

Most government policy decisions have winners and losers. What are the effects of raising the minimum wage? It is more complex than simply producers lose and workers gain. Who are the winners and who are the losers, and what exactly do they win and lose? To what extent does the policy change achieve its goals?

Exercise:**Problem:**

Agricultural price supports result in governments holding large inventories of agricultural products. Why do you think the government cannot simply give the products away to poor people?

Exercise:**Problem:**

Can you propose a policy that would induce the market to supply more rental housing units?

Problems**Exercise:**

Problem:

A low-income country decides to set a price ceiling on bread so it can make sure that bread is affordable to the poor. The conditions of demand and supply are given in [\[link\]](#). What are the equilibrium price and equilibrium quantity before the price ceiling? What will the excess demand or the shortage (that is, quantity demanded minus quantity supplied) be if the price ceiling is set at \$2.40? At \$2.00? At \$3.60?

Price	Qd	Qs
\$1.60	9,000	5,000
\$2.00	8,500	5,500
\$2.40	8,000	6,400
\$2.80	7,500	7,500
\$3.20	7,000	9,000
\$3.60	6,500	11,000
\$4.00	6,000	15,000

Glossary

price ceiling

a legal maximum price

price control

government laws to regulate prices instead of letting market forces determine prices

price floor

a legal minimum price

total surplus

see social surplus

Demand, Supply, and Efficiency

By the end of this section, you will be able to:

- Contrast consumer surplus, producer surplus, and social surplus
- Explain why price floors and price ceilings can be inefficient
- Analyze demand and supply as a social adjustment mechanism

The familiar demand and supply diagram holds within it the concept of economic efficiency. One typical way that economists define efficiency is when it is impossible to improve the situation of one party without imposing a cost on another. Conversely, if a situation is inefficient, it becomes possible to benefit at least one party without imposing costs on others.

Efficiency in the demand and supply model has the same basic meaning: The economy is getting as much benefit as possible from its scarce resources and all the possible gains from trade have been achieved. In other words, the optimal amount of each good and service is being produced and consumed.

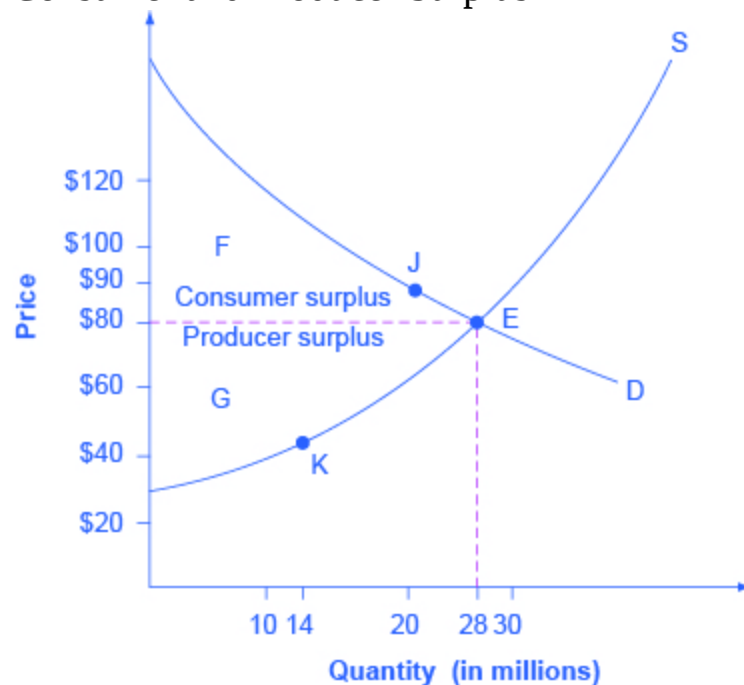
Consumer Surplus, Producer Surplus, Social Surplus

Consider a market for tablet computers, as shown in [\[link\]](#). The equilibrium price is \$80 and the equilibrium quantity is 28 million. To see the benefits to consumers, look at the segment of the demand curve above the equilibrium point and to the left. This portion of the demand curve shows that at least some demanders would have been willing to pay more than \$80 for a tablet.

For example, point J shows that if the price was \$90, 20 million tablets would be sold. Those consumers who would have been willing to pay \$90 for a tablet based on the utility they expect to receive from it, but who were able to pay the equilibrium price of \$80, clearly received a benefit beyond what they had to pay for. Remember, the demand curve traces consumers' willingness to pay for different quantities. The amount that individuals would have been willing to pay, minus the amount that they actually paid, is

called **consumer surplus**. Consumer surplus is the area labeled F—that is, the area above the market price and below the demand curve.

Consumer and Producer Surplus



The somewhat triangular area labeled by F shows the area of consumer surplus, which shows that the equilibrium price in the market was less than what many of the consumers were willing to pay. Point J on the demand curve shows that, even at the price of \$90, consumers would have been willing to purchase a quantity of 20 million. The somewhat triangular area labeled by G shows the area of producer surplus, which shows that the equilibrium price received in the market was more than what many of the producers were willing to accept for their products. For example, point K on the supply curve shows that at a price of \$45, firms would have been willing to supply a quantity of 14 million.

The supply curve shows the quantity that firms are willing to supply at each price. For example, point K in [\[link\]](#) illustrates that, at \$45, firms would still have been willing to supply a quantity of 14 million. Those producers who would have been willing to supply the tablets at \$45, but who were instead able to charge the equilibrium price of \$80, clearly received an extra benefit beyond what they required to supply the product. The amount that a seller is paid for a good minus the seller's actual cost is called **producer surplus**. In [\[link\]](#), producer surplus is the area labeled G—that is, the area between the market price and the segment of the supply curve below the equilibrium.

The sum of consumer surplus and producer surplus is **social surplus**, also referred to as **economic surplus** or **total surplus**. In [\[link\]](#), social surplus would be shown as the area $F + G$. Social surplus is larger at equilibrium quantity and price than it would be at any other quantity. This demonstrates the economic efficiency of the market equilibrium. In addition, at the efficient level of output, it is impossible to produce greater consumer surplus without reducing producer surplus, and it is impossible to produce greater producer surplus without reducing consumer surplus.

Inefficiency of Price Floors and Price Ceilings

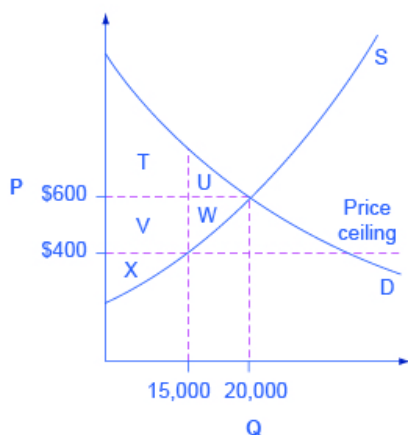
The imposition of a price floor or a price ceiling will prevent a market from adjusting to its equilibrium price and quantity, and thus will create an inefficient outcome. But there is an additional twist here. Along with creating inefficiency, price floors and ceilings will also transfer some consumer surplus to producers, or some producer surplus to consumers.

Imagine that several firms develop a promising but expensive new drug for treating back pain. If this therapy is left to the market, the equilibrium price will be \$600 per month and 20,000 people will use the drug, as shown in [\[link\]](#) (a). The original level of consumer surplus is $T + U$ and producer surplus is $V + W + X$. However, the government decides to impose a price ceiling of \$400 to make the drug more affordable. At this price ceiling, firms in the market now produce only 15,000.

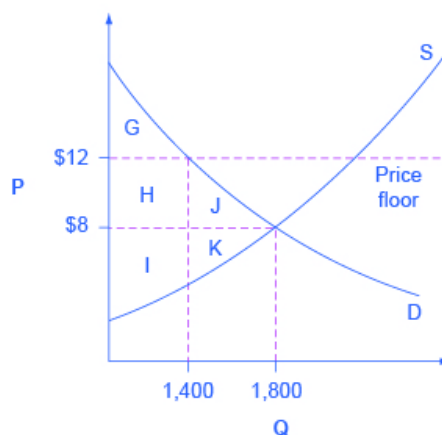
As a result, two changes occur. First, an inefficient outcome occurs and the total surplus of society is reduced. The loss in social surplus that occurs when the economy produces at an inefficient quantity is called **deadweight loss**. In a very real sense, it is like money thrown away that benefits no one. In [\[link\]](#) (a), the deadweight loss is the area U + W. When deadweight loss exists, it is possible for both consumer and producer surplus to be higher, in this case because the price control is blocking some suppliers and demanders from transactions they would both be willing to make.

A second change from the price ceiling is that some of the producer surplus is transferred to consumers. After the price ceiling is imposed, the new consumer surplus is T + V, while the new producer surplus is X. In other words, the price ceiling transfers the area of surplus (V) from producers to consumers. Note that the gain to consumers is less than the loss to producers, which is just another way of seeing the deadweight loss.

Efficiency and Price Floors and Ceilings



(a) Reduced social surplus from a price ceiling



(b) Reduced social surplus from a price floor

- (a) The original equilibrium price is \$600 with a quantity of 20,000. Consumer surplus is T + U, and producer surplus is V + W + X. A price ceiling is imposed at \$400, so firms in the market now produce only a quantity of 15,000. As a result, the new consumer surplus is T + V, while the new producer surplus is X. (b) The original equilibrium is \$8 at a quantity of 1,800. Consumer surplus is G + H + J, and producer surplus is I + K. A price floor is imposed at \$12, which means that quantity demanded falls to 1,400. As a result, the new consumer surplus is G, and the new producer surplus is H + I.

[\[link\]](#) (b) shows a price floor example using a string of struggling movie theaters, all in the same city. The current equilibrium is \$8 per movie ticket, with 1,800 people attending movies. The original consumer surplus is $G + H + J$, and producer surplus is $I + K$. The city government is worried that movie theaters will go out of business, reducing the entertainment options available to citizens, so it decides to impose a price floor of \$12 per ticket. As a result, the quantity demanded of movie tickets falls to 1,400. The new consumer surplus is G , and the new producer surplus is $H + I$. In effect, the price floor causes the area H to be transferred from consumer to producer surplus, but also causes a deadweight loss of $J + K$.

This analysis shows that a price ceiling, like a law establishing rent controls, will transfer some producer surplus to consumers—which helps to explain why consumers often favor them. Conversely, a price floor like a guarantee that farmers will receive a certain price for their crops will transfer some consumer surplus to producers, which explains why producers often favor them. However, both price floors and price ceilings block some transactions that buyers and sellers would have been willing to make, and creates deadweight loss. Removing such barriers, so that prices and quantities can adjust to their equilibrium level, will increase the economy's social surplus.

Demand and Supply as a Social Adjustment Mechanism

The demand and supply model emphasizes that prices are not set only by demand or only by supply, but by the interaction between the two. In 1890, the famous economist Alfred Marshall wrote that asking whether supply or demand determined a price was like arguing “whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper.” The answer is that both blades of the demand and supply scissors are always involved.

The adjustments of equilibrium price and quantity in a market-oriented economy often occur without much government direction or oversight. If the coffee crop in Brazil suffers a terrible frost, then the supply curve of

coffee shifts to the left and the price of coffee rises. Some people—call them the coffee addicts—continue to drink coffee and pay the higher price. Others switch to tea or soft drinks. No government commission is needed to figure out how to adjust coffee prices, which companies will be allowed to process the remaining supply, which supermarkets in which cities will get how much coffee to sell, or which consumers will ultimately be allowed to drink the brew. Such adjustments in response to price changes happen all the time in a market economy, often so smoothly and rapidly that we barely notice them.

Think for a moment of all the seasonal foods that are available and inexpensive at certain times of the year, like fresh corn in midsummer, but more expensive at other times of the year. People alter their diets and restaurants alter their menus in response to these fluctuations in prices without fuss or fanfare. For both the U.S. economy and the world economy as a whole, markets—that is, demand and supply—are the primary social mechanism for answering the basic questions about what is produced, how it is produced, and for whom it is produced.

Note:**Why Can We Not Get Enough of Organic?**

Organic food is grown without synthetic pesticides, chemical fertilizers or genetically modified seeds. In recent decades, the demand for organic products has increased dramatically. The Organic Trade Association reported sales increased from \$1 billion in 1990 to \$35.1 billion in 2013, more than 90% of which were sales of food products.

Why, then, are organic foods more expensive than their conventional counterparts? The answer is a clear application of the theories of supply and demand. As people have learned more about the harmful effects of chemical fertilizers, growth hormones, pesticides and the like from large-scale factory farming, our tastes and preferences for safer, organic foods have increased. This change in tastes has been reinforced by increases in income, which allow people to purchase pricier products, and has made organic foods more mainstream. This has led to an increased demand for organic foods. Graphically, the demand curve has shifted right, and we

have moved up the supply curve as producers have responded to the higher prices by supplying a greater quantity.

In addition to the movement along the supply curve, we have also had an increase in the number of farmers converting to organic farming over time. This is represented by a shift to the right of the supply curve. Since both demand and supply have shifted to the right, the resulting equilibrium quantity of organic foods is definitely higher, but the price will only fall when the increase in supply is larger than the increase in demand. We may need more time before we see lower prices in organic foods. Since the production costs of these foods may remain higher than conventional farming, because organic fertilizers and pest management techniques are more expensive, they may never fully catch up with the lower prices of non-organic foods.

As a final, specific example: The Environmental Working Group's "Dirty Dozen" list of fruits and vegetables, which test high for pesticide residue even after washing, was released in April 2013. The inclusion of strawberries on the list has led to an increase in demand for organic strawberries, resulting in both a higher equilibrium price and quantity of sales.

Consumer surplus is the gap between the price that consumers are willing to pay, based on their preferences, and the market equilibrium price. Producer surplus is the gap between the price for which producers are willing to sell a product, based on their costs, and the market equilibrium price. Social surplus is the sum of consumer surplus and producer surplus. Total surplus is larger at the equilibrium quantity and price than it will be at any other quantity and price. Deadweight loss is loss in total surplus that occurs when the economy produces at an inefficient quantity.

Exercise:

Problem:

Does a price ceiling increase or decrease the number of transactions in a market? Why? What about a price floor?

Solution:

Assuming that people obey the price ceiling, the market price will be below equilibrium, which means that Q_d will be more than Q_s . Buyers can only buy what is offered for sale, so the number of transactions will fall to Q_s . This is easy to see graphically. By analogous reasoning, with a price floor the market price will be above the equilibrium price, so Q_d will be less than Q_s . Since the limit on transactions here is demand, the number of transactions will fall to Q_d . Note that because both price floors and price ceilings reduce the number of transactions, social surplus is less.

Exercise:

Problem:

If a price floor benefits producers, why does a price floor reduce social surplus?

Solution:

Because the losses to consumers are greater than the benefits to producers, so the net effect is negative. Since the lost consumer surplus is greater than the additional producer surplus, social surplus falls.

Exercise:

Problem:

What is consumer surplus? How is it illustrated on a demand and supply diagram?

Exercise:

Problem:

What is producer surplus? How is it illustrated on a demand and supply diagram?

Exercise:

Problem:

What is total surplus? How is it illustrated on a demand and supply diagram?

Exercise:

Problem:

What is the relationship between total surplus and economic efficiency?

Exercise:

Problem: What is deadweight loss?

Exercise:

Problem:

What term would an economist use to describe what happens when a shopper gets a “good deal” on a product?

Exercise:

Problem: Explain why voluntary transactions improve social welfare.

Exercise:

Problem:

Why would a free market never operate at a quantity greater than the equilibrium quantity? *Hint:* What would be required for a transaction to occur at that quantity?

Glossary

consumer surplus

the extra benefit consumers receive from buying a good or service, measured by what the individuals would have been willing to pay

minus the amount that they actually paid

deadweight loss

the loss in social surplus that occurs when a market produces an inefficient quantity

economic surplus

see social surplus

producer surplus

the extra benefit producers receive from selling a good or service, measured by the price the producer actually received minus the price the producer would have been willing to accept

social surplus

the sum of consumer surplus and producer surplus

Negative Externalities class="introduction" Environmental Debate

Across the
country,
countless people
have protested,
even risking
arrest, against
the Keystone
XL Pipeline.

(Credit:
modification of
image by
“NoKXL”/Flick
r Creative
Commons)



Note:
Keystone XL

You might have heard about Keystone XL in the news. It is a pipeline system designed to bring oil from Canada to the refineries near the Gulf of Mexico, as well as to boost crude oil production in the United States. While a private company, TransCanada, will own the pipeline, U.S. government approval is required because of its size and location. The pipeline is being built in four phases, with the first two currently in operation, bringing oil from Alberta, Canada, east across Canada, south through the United States into Nebraska and Oklahoma, and northeast again to Illinois. The third and fourth phases of the project, known as Keystone XL, would create a pipeline southeast from Alberta straight to Nebraska, and then from Oklahoma to the Gulf of Mexico.

Sounds like a great idea, right? A pipeline that would move much needed crude oil to the Gulf refineries would increase oil production for manufacturing needs, reduce price pressure at the gas pump, and increase overall economic growth. Supporters argue that the pipeline is one of the safest pipelines built yet, and would reduce America's dependence on politically vulnerable Middle Eastern oil imports.

Not so fast, say its critics. The Keystone XL would be constructed over an enormous aquifer (one of the largest in the world) in the Midwest, and through an environmentally fragile area in Nebraska, causing great concern among environmentalists about possible destruction to the natural surroundings. They argue that leaks could taint valuable water sources and construction of the pipeline could disrupt and even harm indigenous species. Environmentalist groups have fought government approval of the proposed construction of the pipeline, and as of press time the pipeline projects remain stalled.

Of course, environmental concerns matter when discussing issues related to economic growth. But how much should they factor in? In the case of the pipeline, how do we know how much damage it would cause when we do not know how to put a value on the environment? Would the benefits of the pipeline outweigh the opportunity cost? The issue of how to balance economic progress with unintended effects on our planet is the subject of this chapter.

Note:

Introduction to Environmental Protection and Negative Externalities

In this chapter, you will learn about:

- The Economics of Pollution
- Command-and-Control Regulation
- Market-Oriented Environmental Tools
- The Benefits and Costs of U.S. Environmental Laws
- International Environmental Issues
- The Tradeoff between Economic Output and Environmental Protection

In 1969, the Cuyahoga River in Ohio was so polluted that it spontaneously burst into flame. Air pollution was so bad at that time that Chattanooga, Tennessee was a city where, as an article from Sports Illustrated put it: “the death rate from tuberculosis was double that of the rest of Tennessee and triple that of the rest of the United States, a city in which the filth in the air was so bad it melted nylon stockings off women’s legs, in which executives kept supplies of clean white shirts in their offices so they could change when a shirt became too gray to be presentable, in which headlights were turned on at high noon because the sun was eclipsed by the gunk in the sky.”

The problem of pollution arises for every economy in the world, whether high-income or low-income, and whether market-oriented or command-oriented. Every country needs to strike some balance between production and environmental quality. This chapter begins by discussing how firms may fail to take certain social costs, like pollution, into their planning if they do not need to pay these costs. Traditionally, policies for environmental protection have focused on governmental limits on how much of each pollutant could be emitted. While this approach has had some success, economists have suggested a range of more flexible, market-oriented policies that reduce pollution at a lower cost. We will consider both approaches, but first let’s see how economists frame and analyze these issues.

The Economics of Pollution

By the end of this section, you will be able to:

- Explain and give examples of positive and negative externalities
- Identify equilibrium price and quantity
- Evaluate how firms can contribute to market failure

From 1970 to 2012, the U.S. population increased by one-third and the size of the U.S. economy more than doubled. Since the 1970s, however, the United States, using a variety of anti-pollution policies, has made genuine progress against a number of pollutants. [\[link\]](#) lists the change in carbon dioxide emissions by users of energy (from residential to industrial) according to the U.S. Energy Information Administration (EIA). The table shows that emissions of certain key air pollutants declined substantially from 2007 to 2012; they dropped 730 million metric tons (MMT) a year—a 12% reduction. This seems to indicate that progress has been made in the United States in reducing overall carbon dioxide emissions, which cause greenhouse gases.

	Primary Fossil Fuels			Purchased Electric Power	Total Primary Fossil Fuels
End-use Sector	Coal	Petroleum	Natural Gas		
Residential	(0)	(14)	(31)	(134)	(179)
Commercial	(2)	(2)	(7)	(126)	(136)
Industrial	(40)	(62)	31	(118)	(191)
Transportation	0	(228)	5	(1)	(224)
Power	(464)	(36)	(122)	-	-
Change 2007–2012	(508)	(342)	121	(378)	(730)

U.S. Carbon Dioxide (CO₂) Emissions from Fossil Fuels Consumed 2007–2012, Million Metric Tons (MMT) per Year (Source: EIA Monthly Energy Review)

about 10% (about 10%) per year (Source: Environmental Energy Review)

Despite the gradual reduction in emissions from fossil fuels, many important environmental issues remain. Along with the still high levels of air and water pollution, other issues include hazardous waste disposal, destruction of wetlands and other wildlife habitats, and the impact on human health from pollution.

Externalities

Private markets, such as the cell phone industry, offer an efficient way to put buyers and sellers together and determine what goods are produced, how they are produced, and who gets them. The principle that voluntary exchange benefits both buyers and sellers is a fundamental building block of the economic way of thinking. But what happens when a voluntary exchange affects a third party who is neither the buyer nor the seller?

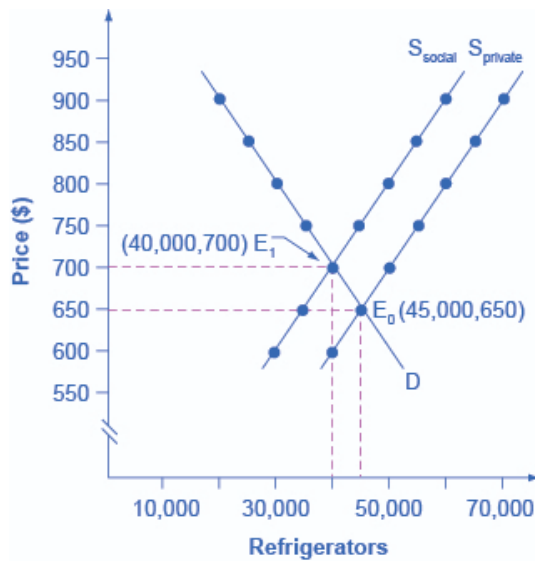
As an example, consider a concert producer who wants to build an outdoor arena that will host country music concerts a half-mile from your neighborhood. You will be able to hear these outdoor concerts while sitting on your back porch—or perhaps even in your dining room. In this case, the sellers and buyers of concert tickets may both be quite satisfied with their voluntary exchange, but you have no voice in their market transaction. The effect of a market exchange on a third party who is outside or “external” to the exchange is called an **externality**. Because externalities that occur in market transactions affect other parties beyond those involved, they are sometimes called **spillovers**.

Externalities can be negative or positive. If you hate country music, then having it waft into your house every night would be a **negative externality**. If you love country music, then what amounts to a series of free concerts would be a **positive externality**.

Pollution as a Negative Externality

Pollution is a negative externality. Economists illustrate the **social costs** of production with a demand and supply diagram. The social costs include the private costs of production incurred by the company and the external costs of pollution that are passed on to society. [\[link\]](#) shows the demand and supply for manufacturing refrigerators. The demand curve (D) shows the quantity demanded at each price. The supply curve (S_{private}) shows the quantity of refrigerators supplied by all the firms at each price if they are taking only their private costs into account and they are allowed to emit pollution at zero cost. The market equilibrium (E_0), where quantity supplied and quantity demanded are equal, is at a price of \$650 and a quantity of 45,000. This information is also reflected in the first three columns of [\[link\]](#).

Taking Social Costs into Account: A Supply Shift



If the firm takes only its own costs of production into account, then its supply curve will be S_{private} , and the market equilibrium will occur at E_0 . Accounting for additional external costs of \$100 for every unit produced, the firm's supply curve will be S_{social} . The new equilibrium will occur at E_1 .

Price	Quantity Demanded	Quantity Supplied before Considering Pollution Cost	Quantity Supplied after Considering Pollution Cost
\$600	50,000	40,000	30,000
\$650	45,000	45,000	35,000
\$700	40,000	50,000	40,000
\$750	35,000	55,000	45,000

Price	Quantity Demanded	Quantity Supplied before Considering Pollution Cost	Quantity Supplied after Considering Pollution Cost
\$800	30,000	60,000	50,000
\$850	25,000	65,000	55,000
\$900	20,000	70,000	60,000

A Supply Shift Caused by Pollution Costs

However, as a by-product of the metals, plastics, chemicals and energy that are used in manufacturing refrigerators, some pollution is created. Let's say that, if these pollutants were emitted into the air and water, they would create costs of \$100 per refrigerator produced. These costs might occur because of injuries to human health, property values, wildlife habitat, reduction of recreation possibilities, or because of other negative impacts. In a market with no anti-pollution restrictions, firms can dispose of certain wastes absolutely free. Now imagine that firms which produce refrigerators must factor in these external costs of pollution—that is, the firms have to consider not only the costs of labor and materials needed to make a refrigerator, but also the broader costs to society of injuries to health and other values caused by pollution. If the firm is required to pay \$100 for the **additional external costs** of pollution each time it produces a refrigerator, production becomes more costly and the entire supply curve shifts up by \$100.

As illustrated in the fourth column of [\[link\]](#) and in [\[link\]](#), the firm will need to receive a price of \$700 per refrigerator and produce a quantity of 40,000—and the firm's new supply curve will be S_{social} . The new equilibrium will occur at E_1 , taking the additional external costs of pollution into account results in a higher price, a lower quantity of production, and a lower quantity of pollution. The following Work It Out feature will walk you through an example, this time with musical accompaniment.

Note:

Identifying the Equilibrium Price and Quantity

[\[link\]](#) shows the supply and demand conditions for a firm that will play trumpets on the streets when requested. Output is measured as the number of songs played.

Price	Quantity Demanded	Quantity Supplied without paying the costs of the externality	Quantity Supplied after paying the costs of the externality
\$20	0	10	8
\$18	1	9	7
\$15	2.5	7.5	5.5
\$12	4	6	4
\$10	5	5	3
\$5	7.5	2.5	0.5

Supply and Demand Conditions for a Trumpet-Playing Firm

Step 1. Determine the negative externality in this situation. To do this, you must think about the situation described and consider all parties that might be impacted. A negative externality might be the increase in noise pollution in the area where the firm is playing.

Step 2. Identify the equilibrium price and quantity when only private costs are taken into account, and then when social costs are taken into account. Remember that equilibrium is where the quantity demanded is equal to the quantity supplied.

Step 3. Look down the columns to where the quantity demanded (the second column) is equal to the “quantity supplied without paying the costs of the externality” (the third column). Then refer to the first column of that row to determine the equilibrium price. In this case, the equilibrium price and quantity when only private costs are taken into account would be at a price of \$10 and a quantity of five.

Step 4. Identify the equilibrium price and quantity when the additional external costs are taken into account. Look down the columns of quantity demanded (the second column) and the “quantity supplied after paying the costs of the externality” (the fourth column) then refer to the first column of that row to determine the equilibrium price. In this case, the equilibrium will be at a price of \$12 and a quantity of four.

Step 5. Consider how taking the externality into account affects the equilibrium price and quantity. Do this by comparing the two equilibrium situations. If the firm is forced to pay its additional external costs, then production of trumpet songs becomes more costly, and the supply curve will shift up.

Remember that the supply curve is based on choices about production that firms make while looking at their marginal costs, while the demand curve is based on the benefits that individuals perceive while maximizing utility. If no externalities existed, private costs

would be the same as the costs to society as a whole, and private benefits would be the same as the benefits to society as a whole. Thus, if no externalities existed, the interaction of demand and supply will coordinate social costs and benefits.

However, when the externality of pollution exists, the supply curve no longer represents all social costs. Because externalities represent a case where markets no longer consider all social costs, but only some of them, economists commonly refer to externalities as an example of **market failure**. When there is market failure, the private market fails to achieve efficient output, because either firms do not account for all costs incurred in the production of output and/or consumers do not account for all benefits obtained (a positive externality). In the case of pollution, at the market output, social costs of production exceed social benefits to consumers, and the market produces too much of the product.

We can see a general lesson here. If firms were required to pay the social costs of pollution, they would create less pollution but produce less of the product and charge a higher price. In the next module, we will explore how governments require firms to take the social costs of pollution into account.

Key Concepts and Summary

Economic production can cause environmental damage. This tradeoff arises for all countries, whether high-income or low-income, and whether their economies are market-oriented or command-oriented.

An externality occurs when an exchange between a buyer and seller has an impact on a third party who is not part of the exchange. An externality, which is sometimes also called a spillover, can have a negative or a positive impact on the third party. If those parties imposing a negative externality on others had to take the broader social cost of their behavior into account, they would have an incentive to reduce the production of whatever is causing the negative externality. In the case of a positive externality, the third party is obtaining benefits from the exchange between a buyer and a seller, but they are not paying for these benefits. If this is the case, then markets would tend to under produce output because suppliers are not aware of the additional demand from others. If the parties that are generating benefits to others would be somehow compensated for these external benefits, they would have an incentive to increase production of whatever is causing the positive externality.

Self-Check Questions

Exercise:

Problem:

Identify the following situations as an example of a negative or a positive externality:

- a. You are a birder (bird watcher), and your neighbor has put up several birdhouses in the yard as well as planting trees and flowers that attract birds.
 - b. Your neighbor paints his house a hideous color.
 - c. Investments in private education raise your country's standard of living.
 - d. Trash dumped upstream flows downstream right past your home.
 - e. Your roommate is a smoker, but you are a nonsmoker.
-

Solution:

- a. positive externality
- b. negative externality
- c. positive externality
- d. negative externality
- e. negative externality

Exercise:

Problem:

Identify whether the market supply curve will shift right or left or will stay the same for the following:

- a. Firms in an industry are required to pay a fine for their emissions of carbon dioxide.
 - b. Companies are sued for polluting the water in a river.
 - c. Power plants in a specific city are not required to address the impact of their emissions on the quality of air.
 - d. Companies that use fracking to remove oil and gas from rock are required to clean up the damage.
-

Solution:

- a. supply shifts left
- b. supply shifts left
- c. supply stays the same
- d. supply shifts left

Exercise:

Problem:

For each of your answers to [\[link\]](#), will equilibrium price rise or fall or stay the same?

Solution:

- a. price will rise
- b. price will rise
- c. price stays the same
- d. price will rise.

Exercise:

Problem:

The supply and demand conditions for a manufacturing firm are given in [\[link\]](#). The third column represents a supply curve without taking the social cost of pollution into account. The fourth column represents the supply curve when the firm is required to take the social cost of pollution into account. Identify the equilibrium before the social cost of production is included and after the social cost of production is included.

Price	Quantity Demanded	Quantity Supplied without paying the cost of the pollution	Quantity Supplied after paying the cost of the pollution
\$10	450	400	250
\$15	440	440	290
\$20	430	480	330
\$25	420	520	370
\$30	410	560	410

Solution:

The original equilibrium (before the external social cost of pollution is taken into account) is where the private supply curve crosses the demand curve. This original equilibrium is at a price of \$15 and a quantity of 440. After taking into account the additional external cost of pollution, the production becomes more costly, and the supply curve shifts up. The new equilibrium will be at a price of \$30 and a quantity of 410.

Review Questions

Exercise:

Problem: What is an externality?

Exercise:

Problem:

Give an example of a positive externality and an example of a negative externality.

Exercise:

Problem: What is the difference between private costs and social costs?

Exercise:

Problem:

In a market without environmental regulations, will the supply curve for a firm take into account private costs, external costs, both, or neither? Explain.

Critical Thinking Questions

Exercise:

Problem:

Suppose you want to put a dollar value on the external costs of carbon emissions from a power plant. What information or data would you obtain to measure the external [not social] cost?

Problems

Exercise:

Problem:

Show the market for cigarettes in equilibrium, assuming that there are no laws banning smoking in public. Label the equilibrium private market price and quantity as P_m and Q_m . Add whatever is needed to the model to show the impact of the negative externality from second-hand smoking. (Hint: In this case it is the consumers, not the sellers, who are creating the negative externality.) Label the social optimal output and price as P_e and Q_e . On the graph, shade in the deadweight loss at the market output.

Exercise:

Problem:

Refer to [\[link\]](#). The externality created by the production of refrigerators was \$100. However, once both the private and additional external costs were taken into consideration, the market price increased by only \$50. If the external costs were \$100 why did the price only increase by \$50 when all costs were taken into account?

Exercise:**Problem:**

[\[link\]](#), shows the supply and demand conditions for a firm that will play trumpets on the streets when requested. Q_{s1} is the quantity supplied without social costs. Q_{s2} is the quantity supplied with social costs. What is the negative externality in this situation? Identify the equilibrium price and quantity when only private costs are taken into account, and then when social costs are taken into account. How does taking the externality into account affect the equilibrium price and quantity?

P	Qd	Q _{s1}	Q _{s2}
\$20	0	10	8
\$18	1	9	7
\$15	2.5	7.5	5.5
\$12	4	6	4
\$10	5	5	3
\$5	7.5	2.5	0.5

References

Johnson, Oscar William. "Back on Track: Earth Day Success Story; The Chattanooga Choo-Choo No Longer Spews Foul Air." *Sports Illustrated*. April 30, 1990.
<http://www.si.com/vault/1990/04/30/121923/back-on-track-earth-day-success-story-the-chattanooga-choo-choo-no-longer-spews-foul-air>.

U.S. Energy Information Administration. "Total Energy: Monthly Energy Review." *U.S. Department of Energy*. Accessed December 19, 2013.
<http://www.eia.gov/totalenergy/data/monthly/>.

Glossary

additional external cost

additional costs incurred by third parties outside the production process when a unit of output is produced

externality

a market exchange that affects a third party who is outside or "external" to the exchange; sometimes called a "spillover"

market failure

When the market on its own does not allocate resources efficiently in a way that balances social costs and benefits; externalities are one example of a market failure

negative externality

a situation where a third party, outside the transaction, suffers from a market transaction by others

positive externality

a situation where a third party, outside the transaction, benefits from a market transaction by others

social costs

costs that include both the private costs incurred by firms and also additional costs incurred by third parties outside the production process, like costs of pollution

spillover

see externality

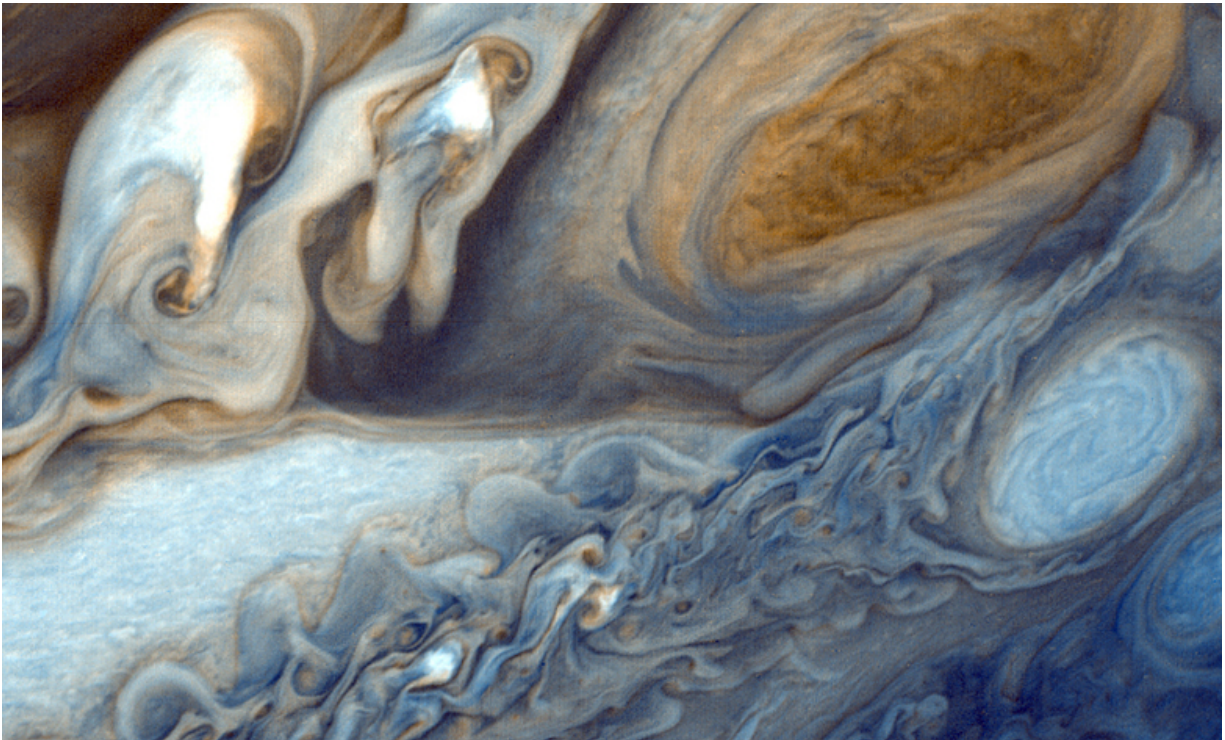
Introduction to Positive Externalities and Public Goods

class="introduction"

View from Voyager I

Launched by
NASA on
September 5,
1977,
Voyager 1's
primary
mission was
to provide
detailed
images of
Jupiter,
Saturn, and
their moons.
It took this
photograph
of Jupiter on
its journey.
In August of
2012,
Voyager I
entered
interstellar
space—the
first human-
made object
to do so—
and it is
expected to
send data
and images
back to earth
until 2025.
Such a

technological feat has a lot to do with economic principles.
(Credit: modification of work by NASA/JPL)



Note:

The Benefits of Voyager I Live On

The rapid growth of technology has increased our ability to access and process data, to navigate through a busy city, and to communicate with friends on the other side of the globe. The research and development efforts of citizens, scientists, firms, universities, and governments have

truly revolutionized the modern economy. To get a sense of how far we have come in a short period of time, let's compare one of humankind's greatest achievements to the smartphone most of us have in our coat pocket.

In 1977 the United States launched Voyager I, a spacecraft originally intended to reach Jupiter and Saturn, to send back photographs and other cosmic measurements. Voyager I, however, kept going, and going—past Jupiter and Saturn—right out of our solar system. At the time of its launch, Voyager had some of the most sophisticated computing processing power NASA could engineer (8,000 instructions per second), but by the time it left the solar system (in 2012, actually) we Earthlings were using handheld devices that could process 14 billion instructions per second.

Still, the technology of today is a spillover product of the incredible feats accomplished by NASA thirty years ago. NASA research, for instance, is responsible for the kidney dialysis and mammogram machines that we use today. Research in new technologies not only produces private benefits to the investing firm, or in this case to NASA, but it also creates benefits for the broader society. In this way, new knowledge often becomes what economists refer to as a public good. This leads us to the topic of this chapter—technology, positive externalities, public goods, and the role of government in the encouragement of innovation and the social benefits that it provides.

Note:**Introduction to Positive Externalities and Public Goods**

In this chapter, you will learn about:

- Why the Private Sector Under Invests in Technologies
- How Governments Can Encourage Innovation
- Public Goods

Can you imagine a world in which you did not own a cellular phone or use Wikipedia? New technology changes how people live and work and what

they buy. Technology includes the invention of new products, new ways of producing goods and services, and even new ways of managing a company more efficiently. Research and development of technology is the difference between horses and automobiles, between candles and electric lights, between fetching water in buckets and indoor plumbing, and between infection and good health from antibiotics.

In December 2009, ABC News compiled a list of some of the technological breakthroughs that have revolutionized consumer products in the past 10 years:

- GPS tracking devices, originally developed by the defense department and available to consumers in 2000, give users up-to-date information on location and time through satellite technology.
- In 2000, Toyota introduced the Prius hybrid car, which greatly improved fuel efficiency.
- Also in 2000, AT&T offered its customers the ability to text on a mobile phone.
- In 2001, Wikipedia launched a user-generated encyclopedia on the Web.
- Even though Napster died in 2001, the company launched music downloading and file sharing, which revolutionized how consumers get their music and videos.
- Friendster kicked off the social networking business in 2003, and Twitter and Facebook followed.
- In 2003, the Human Genome project was completed. It helps to fight disease and launch new pharmaceutical innovations.
- Also in 2003, the search engine became a way of life for obtaining information quickly. The search engine companies also became innovators in the digital software that dominates mobile devices.
- In 2006, Nintendo launched Wii and changed the way video games are played. Players can now be drawn into the action and use their bodies to respond rather than a handheld device.
- Apple introduced the iPhone in 2007 and launched an entire smartphone industry. In 2015, cell phones now recognize human voices via artificial intelligence.

With all new technologies, however, there are new challenges. This chapter deals with some of these issues: Will private companies be willing to invest in new technology? In what ways does new technology have positive externalities? What motivates inventors? Does government have a role to play in encouraging research and technology? Are there certain types of goods that markets fail to provide efficiently, and that only government can produce? What happens when consumption or production of a product creates positive externalities? Why is it unsurprising when a common resource, like marine fisheries, is overused?

How Governments Can Encourage Innovation

By the end of this section, you will be able to:

- Explain the effects of intellectual property rights on social and private rates of return.
- Identify three U.S. Government policies and explain how they encourage innovation

A number of different government policies can increase the incentives to innovate, including: guaranteeing intellectual property rights, government assistance with the costs of research and development, and cooperative research ventures between universities and companies.

Intellectual Property Rights

One way to increase new technology is to guarantee the innovator an exclusive right to that new product or process. **Intellectual property** rights include patents, which give the inventor the exclusive legal right to make, use, or sell the invention for a limited time, and copyright laws, which give the author an exclusive legal right over works of literature, music, film/video, and pictures. For example, if a pharmaceutical firm has a patent on a new drug, then no other firm can manufacture or sell that drug for 21 years, unless the firm with the patent grants permission. Without a patent, the pharmaceutical firm would have to face competition for any successful products, and could earn no more than a normal rate of profit. With a patent, a firm is able to earn monopoly profits on its product for a period of time—which offers an incentive for research and development. In general, how long can “a period of time” be? The Clear It Up discusses patent and copyright protection timeframes for some works you might know.

Note:

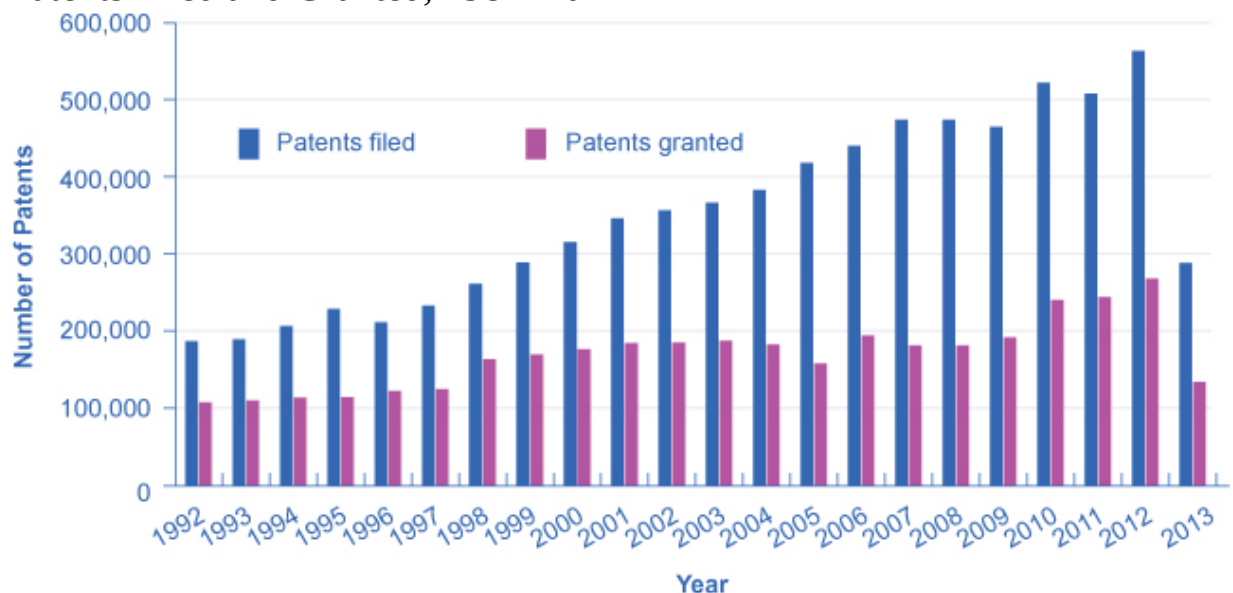
How long is Mickey Mouse protected from being copied?

All patents and copyrights are scheduled to end someday. In 2003, copyright protection for Mickey Mouse was scheduled to run out. Once the copyright had expired, anyone would be able to copy Mickey Mouse

cartoons or draw and sell new ones. In 1998, however, Congress passed the Sonny Bono Copyright Term Extension Act. For copyrights owned by companies or other entities, it increased or extended the copyright from 75 years to 95 years after publication. For copyrights owned by individuals, it increased or extended the copyright coverage from 50 years to 70 years after death. Along with protecting Mickey for another 20 years, the copyright extension affected about 400,000 books, movies, and songs.

[\[link\]](#) illustrates how the total number of patent applications filed with the U.S. Patent and Trademark Office, as well as the total number of patents granted, surged in the mid-1990s with the invention of the internet, and is still going strong today.

Patents Filed and Granted, 1981–2012



The number of applications filed for patents increased substantially beginning in the 1990s, due in part to the invention of the internet, which has led to many other inventions and to the 1998 Copyright

Term Extension Act. (Source:

http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm)

While patents provide an incentive to innovate by protecting the innovator, they are not perfect. For example:

- In countries that already have patents, economic studies show that inventors receive only one-third to one-half of the total economic value of their inventions.
- In a fast-moving high-technology industry like biotechnology or semiconductor design, patents may be almost irrelevant because technology is advancing so quickly.
- Not every new idea can be protected with a patent or a copyright—for example, a new way of organizing a factory or a new way of training employees.
- Patents may sometimes cover too much or be granted too easily. In the early 1970s, Xerox had received over 1,700 patents on various elements of the photocopy machine. Every time Xerox improved the photocopier, it received a patent on the improvement.
- The 21-year time period for a patent is somewhat arbitrary. Ideally, a patent should cover a long enough period of time for the inventor to earn a good return, but not so long that it allows the inventor to charge a monopoly price permanently.

Because patents are imperfect and do not apply well to all situations, alternative methods of improving the rate of return for inventors of new technology are desirable. The following sections describe some of these possible alternative policies.

Policy #1: Government Spending on Research and Development

If the private sector does not have sufficient incentive to carry out research and development, one possibility is for the government to fund such work directly. Government spending can provide direct financial support for research and development (R&D) conducted at colleges and universities, nonprofit research entities, and sometimes by private firms, as well as at government-run laboratories. While government spending on research and development produces technology that is broadly available for firms to use,

it costs taxpayers money and can sometimes be directed more for political than for scientific or economic reasons.

Note:
Visit the NASA [website](#) and the USDA [website](#) to read about government research that would not take place were it left to firms due to the externalities.

The first column of [\[link\]](#) shows the sources of total U.S. spending on research and development. The second column shows the total dollars of R&D funding by each source. The third column shows that, relative to the total amount of funding, 22.7% comes from the federal government, about 69% of R&D is done by industry, and less than 4% is done by universities and colleges. (The percentages below do not add up to exactly 100% due to rounding.)

Sources of R&D Funding	Amount (\$ billions)	Percent of the Total
Federal government	\$113.1	22.7%
Industry	\$344.9	69.0%
Universities and colleges	\$17.1	3.4%
Nonprofits	\$19.9	4.0%

Sources of R&D Funding	Amount (\$ billions)	Percent of the Total
Nonfederal government	\$4.0	0.8%
<i>Total</i>	<i>\$499</i>	

U.S. Research and Development Expenditures, 2015(Source: <https://www.nsf.gov/statistics/2016/nsf16316/>)

In the 1960s the federal government paid for about two-thirds of the nation's R&D. Over time, the U.S. economy has come to rely much more heavily on industry-funded R&D. The federal government has tried to focus its direct R&D spending on areas where private firms are not as active. One difficulty with direct government support of R&D is that it inevitably involves political decisions about which projects are worthy. The scientific question of whether research is worthwhile can easily become entangled with considerations like the location of the congressional district in which the research funding is spent.

Policy #2: Tax Breaks for Research and Development

A complementary approach to supporting R&D that does not involve the government's close scrutiny of specific projects is to give firms a reduction in taxes depending on how much research and development they do. The federal government refers to this policy as the research and experimentation (R&E) tax credit. According to the Treasury Department: “. . . the R&E Credit is also a cost-effective policy for stimulating additional private sector investment. Most recent studies find that each dollar of foregone tax revenue through the R&E Tax Credit causes firms to invest at least a dollar in R&D, with some studies finding a benefit to cost ratio of 2 or 2.96.”

Note:

Visit this [website](#) for more information on how the R&E Tax Credit encourages investment.

Policy #3 Cooperative Research

State and federal governments support research in a variety of ways. For example, United for Medical Research, a coalition of groups that seek funding for the National Institutes of Health, (which is supported by federal grants), states: “NIH-supported research added \$69 billion to our GDP and supported seven million jobs in 2011 alone.” The United States remains the leading sponsor of medical-related research spending \$117 billion in 2011. Other institutions, such as the National Academy of Scientists and the National Academy of Engineers, receive federal grants for innovative projects. The Agriculture and Food Research Initiative (AFRI) at the United States Department of Agriculture awards federal grants to projects that apply the best science to the most important agricultural problems, from food safety to childhood obesity. Cooperation between government-funded universities, academies, and the private sector can spur product innovation and create whole new industries.

Key Concepts and Summary

Public policy with regard to technology must often strike a balance. For example, patents provide an incentive for inventors, but they should be limited to genuinely new inventions and not extend forever.

Government has a variety of policy tools for increasing the rate of return for new technology and encouraging its development, including: direct government funding of R&D, tax incentives for R&D, protection of intellectual property, and forming cooperative relationships between universities and the private sector.

Self-Check Questions

Exercise:

Problem:

When residents in a neighborhood tidy it and keep it neat, there are a number of positive spillovers: higher property values, less crime, happier residents. What types of government policies can encourage neighborhoods to clean up?

Solution:

Government programs that either pay for neighborhood clean-up directly or that provide reduced tax payments for those who clean up or fix up their own property could be enacted. It is also easy to imagine how a city might allow its businesses to form a group that would pay for and manage neighborhood cleanup.

Exercise:**Problem:**

Education provides both private benefits to those who receive it and broader social benefits for the economy as a whole. Think about the types of policies a government can follow to address the issue of positive spillovers in technology and then suggest a parallel set of policies that governments could follow for addressing positive externalities in education.

Solution:

Government programs that either pay for education directly or that provide loans or reduced tax payments for education could create positive spillovers. A city might allow its businesses to form a group that would coordinate business efforts with schools and local colleges and universities—allowing students to obtain real-world experience in their chosen fields and providing businesses with enthusiastic, trained workers.

Review Questions

Exercise:**Problem:**

Why might private markets tend to provide too few incentives for the development of new technology?

Exercise:**Problem:**

What can government do to encourage the development of new technology?

Critical Thinking Question**Exercise:****Problem:**

Is it inevitable that government must become involved in supporting investments in new technology?

Problem**Exercise:****Problem:**

Assume that the marginal private costs of a firm producing fuel-efficient cars is greater than the marginal social costs. Assume that the marginal private benefits of a firm producing fuel-efficient cars is the same as the marginal social benefits. Discuss one way that the government can try to increase production and sales of fuel efficient cars to the socially desirable amount. *Hint:* the government is trying to affect production through costs, not benefits.

References

United States Department of the Treasury. “Research and Experimentation Tax Credit.” Accessed November 2013. <http://www.treasury.gov/press-center/news/Pages/investing-in-us-competitiveness.aspx>.

U.S. Patent and Trademark Office. 2015. “U.S. Patent Statistics: Calendar Years 1963–2014.” Accessed April 10, 2015. http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.pdf.

United for Medical Research. “Profiles of Prosperity: How NIH-Supported Research Is Fueling Private Sector Growth and Innovation.” Introduction. Accessed January 2014. http://www.unitedformedicalresearch.com/wp-content/uploads/2013/07/UMR_ProspersityReport_071913a.pdf.

Glossary

intellectual property

the body of law including patents, trademarks, copyrights, and trade secret law that protect the right of inventors to produce and sell their inventions

Public Goods

By the end of this section, you will be able to:

- Identify a public good using nonexcludable and nonrivalrous as criteria
- Explain the free rider problem
- Identify several sources of public goods

Even though new technology creates positive externalities so that perhaps one-third or one-half of the social benefit of new inventions spills over to others, the inventor still receives some private return. What about a situation where the positive externalities are so extensive that private firms could not expect to receive any of the social benefit? This kind of good is called a **public good**. Spending on national defense is a good example of a public good. Let's begin by defining the characteristics of a public good and discussing why these characteristics make it difficult for private firms to supply public goods. Then we will see how government may step in to address the issue.

The Definition of a Public Good

Economists have a strict definition of a public good, and it does not necessarily include all goods financed through taxes. To understand the defining characteristics of a public good, first consider an ordinary private good, like a piece of pizza. A piece of pizza can be bought and sold fairly easily because it is a separate and identifiable item. However, public goods are not separate and identifiable in this way.

Instead, public goods have two defining characteristics: they are nonexcludable and nonrivalrous. The first characteristic, that a public good is **nonexcludable**, means that it is costly or impossible to exclude someone from using the good. If Larry buys a private good like a piece of pizza, then he can exclude others, like Lorna, from eating that pizza. However, if national defense is being provided, then it includes everyone. Even if you strongly disagree with America's defense policies or with the level of defense spending, the national defense still protects you. You cannot choose to be unprotected, and national defense cannot protect everyone else and exclude you.

The second main characteristic of a public good, that it is **nonrivalrous**, means that when one person uses the public good, another can also use it. With a private good like pizza, if Max is eating the pizza then Michelle cannot also eat it; that is, the two people are rivals in consumption. With a public good like national defense, Max's consumption of national defense does not reduce the amount left for Michelle, so they are nonrivalrous in this area.

A number of government services are examples of public goods. For instance, it would not be easy to provide fire and police service so that some people in a neighborhood would be protected from the burning and burglary of their property, while others would not be protected at all. Protecting some necessarily means protecting others, too.

Positive externalities and public goods are closely related concepts. Public goods have positive externalities, like police protection or public health funding. Not all goods and services with positive externalities, however, are public goods. Investments in education have huge positive spillovers but can be provided by a private company. Private companies can invest in new inventions such as the Apple iPad and reap profits that may not capture all of the social benefits. Patents can also be described as an attempt to make new inventions into private goods, which are excludable and rivalrous, so that no one but the inventor is allowed to use them during the length of the patent.

The Free Rider Problem of Public Goods

Private companies find it difficult to produce public goods. If a good or service is nonexcludable, like national defense, so that it is impossible or very costly to exclude people from using this good or service, then how can a firm charge people for it?

Note:

Visit this [website](#) to read about a connection between free riders and “bad music.”



When individuals make decisions about buying a public good, a **free rider** problem can arise, in which people have an incentive to let others pay for the public good and then to “free ride” on the purchases of others. The free rider problem can be expressed in terms of the prisoner’s dilemma game, which is discussed as a representation of oligopoly in [Monopolistic Competition and Oligopoly](#). Say that two people are thinking about contributing to a public good: Rachel and Samuel. When either of them contributes to a public good, such as a local fire department, their personal cost of doing so is \$4 and the social benefit of that person’s contribution is \$6. Because society’s benefit of \$6 is greater than the cost of \$4, the investment is a good idea for society as a whole. The problem is that, while Rachel and Samuel pay for the entire cost of their contribution to the public good, they receive only half of the benefit, because the benefit of the public good is divided equally among the members of society. This sets up the prisoner’s dilemma illustrated in [\[link\]](#).

	Samuel (S) Contribute	Samuel (S) Do Not Contribute
Rachel (R) Contribute	R pays \$4, receives \$6, net gain +\$2 S pays \$4, receives \$6, net gain +\$2	R pays \$4, receives \$3, net gain –\$1 S pays \$0, receives \$3, net gain +\$3

Rachel (R) Do Not Contribute	R pays \$0, receives \$3, net gain +\$3 S pays \$4, receives \$3, net gain -\$1	R pays \$0, receives \$0 S pays \$0, receives \$0
---------------------------------	--	--

Contributing to a Public Good as a Prisoner's Dilemma

If neither Rachel nor Samuel contributes to the public good, then there are no costs and no benefits of the public good. Suppose, however, that only Rachel contributes, while Samuel does not. Rachel incurs a cost of \$4, but receives only \$3 of benefit (half of the total \$6 of benefit to society), while Samuel incurs no cost, and yet he also receives \$3 of benefit. In this outcome, Rachel actually loses \$1 while Samuel gains \$3. A similar outcome, albeit with roles reversed, would occur if Samuel had contributed, but Rachel had not. Finally, if both parties contribute, then each incurs a cost of \$4 and each receives \$6 of benefit (half of the total \$12 benefit to society). There is a dilemma with the Prisoner's Dilemma, though. See the Work it Out feature.

Note:

The Problem with the Prisoner's Dilemma

The difficulty with the prisoner's dilemma arises as each person thinks through his or her strategic choices.

Step 1. Rachel reasons in this way: If Samuel does not contribute, then I would be a fool to contribute. However, if Samuel does contribute, then I can come out ahead by not contributing.

Step 2. Either way, I should choose not to contribute, and instead hope that I can be a free rider who uses the public good paid for by Samuel.

Step 3. Samuel reasons the same way about Rachel.

Step 4. When both people reason in that way, the public good never gets built, and there is no movement to the option where everyone cooperates—which is actually best for all parties.

The Role of Government in Paying for Public Goods

The key insight in paying for public goods is to find a way of assuring that everyone will make a contribution and to prevent free riders. For example, if people come together through the political process and agree to pay taxes and make group decisions about the quantity of public goods, they can defeat the free rider problem by requiring, through the law, that everyone contributes.

However, government spending and taxes are not the only way to provide public goods. In some cases, markets can produce public goods. For example, think about radio. It is nonexcludable, since once the radio signal is being broadcast, it would be very difficult to stop someone from receiving it. It is nonrivalrous, since one person listening to the signal does not prevent others from listening as well. Because of these features, it is practically impossible to charge listeners directly for listening to conventional radio broadcasts.

Radio has found a way to collect revenue by selling advertising, which is an indirect way of “charging” listeners by taking up some of their time. Ultimately, consumers who purchase the goods advertised are also paying for the radio service, since the cost of advertising is built into the product cost. In a more recent development, satellite radio companies, such as SiriusXM, charge a regular subscription fee for streaming music without commercials. In this case, however, the product is excludable—only those who pay for the subscription will receive the broadcast.

Some public goods will also have a mixture of public provision at no charge along with fees for some purposes, like a public city park that is free to use, but the government charges a fee for parking your car, for reserving certain picnic grounds, and for food sold at a refreshment stand.

Note:

Read this [article](#) to find out what economists say the government should pay for.



In other cases, social pressures and personal appeals can be used, rather than the force of law, to reduce the number of free riders and to collect resources for the public good. For example, neighbors sometimes form an association to carry out beautification projects or to patrol their area after dark to discourage crime. In low-income countries, where social pressure strongly encourages all farmers to participate, farmers in a region may come together to work on a large irrigation project that will benefit all. Many fundraising efforts, including raising money for local charities and for the endowments of colleges and universities, also can be viewed as an attempt to use social pressure to discourage free riding and to generate the outcome that will produce a public benefit.

Common Resources and the “Tragedy of the Commons”

There are some goods that do not fall neatly into the categories of private good or public good. While it is easy to classify a pizza as a private good and a city park as a public good, what about an item that is nonexcludable and rivalrous, such as the queen conch?

In the Caribbean, the queen conch is a large marine mollusk found in shallow waters of sea grass. These waters are so shallow, and so clear, that a single diver may harvest many conch in a single day. Not only is conch meat a local delicacy and an important part of the local diet, but the large ornate shells are used in art and can be crafted into musical instruments. Because almost anyone with a small boat, snorkel, and mask, can participate in the conch harvest, it is essentially nonexcludable. At the same time, fishing for conch is rivalrous; once a diver catches one conch it cannot be caught by another diver.

Goods that are nonexcludable and rivalrous are called common resources. Because the waters of the Caribbean are open to all conch fishermen, and because any conch that *you* catch is conch that *I* cannot catch, common resources like the conch tend to be overharvested.

The problem of overharvesting common resources is not a new one, but ecologist Garret Hardin put the tag “Tragedy of the Commons” to the problem in a 1968 article in the magazine *Science*. Economists view this as a problem of property rights. Since nobody owns the ocean, or the conch that crawl on the sand beneath it, no one individual has an incentive to protect that resource and responsibly harvest it. To address the issue of overharvesting conch and other marine fisheries, economists typically advocate simple devices like fishing licenses, harvest limits, and shorter fishing seasons. When the population of a species drops to critically low numbers, governments have even banned the harvest until biologists determine that the population has returned to sustainable levels. In fact, such is the case with the conch, the harvesting of which has been effectively banned in the United States since 1986.

Note:

Visit this [website](#) for more on the queen conch industry.



Positive Externalities in Public Health Programs

One of the most remarkable changes in the standard of living in the last several centuries is that people are living longer. Thousands of years ago,

human life expectancy is believed to have been in the range of 20 to 30 years. By 1900, average life expectancy in the United States was 47 years. By 2015, life expectancy is 79 years. Most of the gains in life expectancy in the history of the human race happened in the twentieth century.

The rise in life expectancy seems to stem from three primary factors. First, systems for providing clean water and disposing of human waste helped to prevent the transmission of many diseases. Second, changes in public behavior have advanced health. Early in the twentieth century, for example, people learned the importance of boiling bottles before using them for food storage and baby's milk, washing their hands, and protecting food from flies. More recent behavioral changes include reducing the number of people who smoke tobacco and precautions to limit sexually transmitted diseases. Third, medicine has played a large role. Immunizations for diphtheria, cholera, pertussis, tuberculosis, tetanus, and yellow fever were developed between 1890 and 1930. Penicillin, discovered in 1941, led to a series of other antibiotic drugs for bringing infectious diseases under control. In recent decades, drugs that reduce the risks of high blood pressure have had a dramatic effect in extending lives.

These advances in public health have all been closely linked to positive externalities and public goods. Public health officials taught hygienic practices to mothers in the early 1900s and encouraged less smoking in the late 1900s. Many public sanitation systems and storm sewers were funded by government because they have the key traits of public goods. In the twentieth century, many medical discoveries came out of government or university-funded research. Patents and intellectual property rights provided an additional incentive for private inventors. The reason for requiring immunizations, phrased in economic terms, is that it prevents spillovers of illness to others—as well as helping the person immunized.

Note:

The Benefits of Voyager I Live On

While we applaud the technology spillovers of NASA's space projects, we should also acknowledge that those benefits are not shared equally.

Economists like Tyler Cowen, a professor at George Mason University, are

seeing more and more evidence of a widening gap between those who have access to rapidly improving technology, and those who do not. According to Cowen author of the recent book, *Average Is Over: Powering America Beyond the Age of the Great Stagnation*, this inequality in access to technology and information is going to deepen the inequality in skills, and ultimately, in wages and global standards of living.

Key Concepts and Summary

A public good has two key characteristics: it is nonexcludable and nonrivalrous. Nonexcludable means that it is costly or impossible for one user to exclude others from using the good. Nonrivalrous means that when one person uses the good, it does not prevent others from using it. Markets often have a difficult time producing public goods because free riders will attempt to use the public good without paying for it. The free rider problem can be overcome through measures to assure that users of the public good pay for it. Such measures include government actions, social pressures, and specific situations where markets have discovered a way to collect payments.

Self-Check Questions

Exercise:

Problem:

Which of the following goods or services are nonexcludable?

- a. police protection
- b. streaming music from satellite transmission programs
- c. roads
- d. primary education
- e. cell phone service

Solution:

- a. Once citizens are protected from crime, it is difficult to exclude someone from this protection, so it is nonexcludable.
- b. Some satellite radio services, such as SiriusXM, are sold by subscription fee, so it is excludable.
- c. Once a road is built it is difficult to exclude people, although toll roads can exclude non-payers.
- d. Primary education can be provided by private companies and so it is excludable.
- e. Companies sell cell phone service and exclude those who do not pay.

Exercise:

Problem: Are the following goods nonrivalrous in consumption?

- a. slice of pizza
- b. laptop computer
- c. public radio
- d. ice cream cone

Solution:

- a. Two people cannot enjoy the same slice of pizza at the same time, so private goods, such as a slice of pizza, are rivalrous.
- b. Two people cannot use one laptop at the same time, so they are rivalrous in consumption.
- c. Public radio can be heard by anyone with a radio, so many people can listen at the same time—the good is nonrivalrous.
- d. It is difficult for two people to simultaneously eat an ice cream cone, so it is rivalrous in consumption.

Review Questions

Exercise:

Problem: What are the two key characteristics of public goods?

Exercise:

Problem:

Name two public goods and explain why they are public goods.

Exercise:

Problem: What is the free rider problem?

Exercise:

Problem: Explain why the federal government funds national defense.

Critical Thinking Questions

Exercise:

Problem:

How do public television stations, like PBS, try to overcome the free rider problem?

Exercise:

Problem:

Why is a football game on ESPN a quasi-public good but a game on the NBC, CBS, or ABC is a public good?

Exercise:

Problem:

Provide two examples of goods/services that are classified as private goods/services even though they are provided by a federal government.

Exercise:

Problem:

Radio stations, tornado sirens, light houses, and street lights are all public goods in that all are nonrivalrous and nonexclusionary. Therefore why does the government provide tornado sirens, street lights and light houses but not radio stations (other than PBS stations)?

Problems**Exercise:****Problem:**

Becky and Sarah are sisters who share a room. Their room can easily get messy, and their parents are always telling them to clean it up. Here are the costs and benefits to both Becky and Sarah, of taking the time to clean their room: If both Becky and Sarah clean, they each spends two hours and get a clean room. If Becky decides not to clean and Sarah does all the cleaning, then Sarah spends 10 hours cleaning (Becky spends 0) but Sarah is exhausted. The same would occur for Becky if Sarah decided not to clean—Becky spends 10 hours and becomes exhausted. If both girls decide not to clean, they both have a dirty room.

- a. What is the best outcome for Becky and Sarah? What is the worst outcome? (It would help you to construct a prisoner's dilemma table.)
- b. Unfortunately, we know that the optimal outcome will most likely not happen, and that the worst one will probably be chosen instead. Explain what it is about Becky's and Sarah's reasoning that will lead them both to choose the worst outcome.

References

Cowen, Tyler. *Average Is Over: Powering America Beyond the Age of the Great Stagnation*. Dutton Adult, 2013.

Hardin, Garret. "The Tragedy of the Commons." *Science* 162 (3859): 1243–48 (1968).

Glossary

free rider

those who want others to pay for the public good and then plan to use the good themselves; if many people act as free riders, the public good may never be provided

nonexcludable

when it is costly or impossible to exclude someone from using the good, and thus hard to charge for it

nonrivalrous

even when one person uses the good, others can also use it

public good

good that is nonexcludable and nonrivalrous, and thus is difficult for market producers to sell to individual consumers

Introduction to the Macroeconomic Perspective

class="introduction"

The Great Depression

At times, such as when many people are in need of government assistance, it is easy to tell how the economy is doing.

This photograph shows people lined up during the Great Depression, waiting for relief checks. At

other times, when some are doing well and others are not, it is more difficult to ascertain how the economy of a country is doing.

(Credit: modification of work by the U.S.

Library of Congress/Wikimedia Commons)

**Note:****How is the Economy Doing? How Does One Tell?**

The 1990s were boom years for the U.S. economy. The late 2000s, from 2007 to 2014 were not. What causes the economy to expand or contract? Why do businesses fail when they are making all the right decisions? Why do workers lose their jobs when they are hardworking and productive? Are bad economic times a failure of the market system? Are they a failure of the government? These are all questions of macroeconomics, which we will begin to address in this chapter. We will not be able to answer all of these questions here, but we will start with the basics: How is the economy doing? How can we tell?

The macro economy includes all buying and selling, all production and consumption; everything that goes on in every market in the economy. How can we get a handle on that? The answer begins more than 80 years ago, during the Great Depression. President Franklin D. Roosevelt and his economic advisers knew things were bad—but how could they express and measure just how bad it was? An economist named Simon Kuznets, who later won the Nobel Prize for his work, came up with a way to track what the entire economy is producing. The result—gross domestic product

(GDP)—remains our basic measure of macroeconomic activity. In this chapter, you will learn how GDP is constructed, how it is used, and why it is so important.

Note:

Introduction to the Macroeconomic Perspective

In this chapter, you will learn about:

- Measuring the Size of the Economy: Gross Domestic Product
- Adjusting Nominal Values to Real Values
- Tracking Real GDP over Time
- Comparing GDP among Countries
- How Well GDP Measures the Well-Being of Society

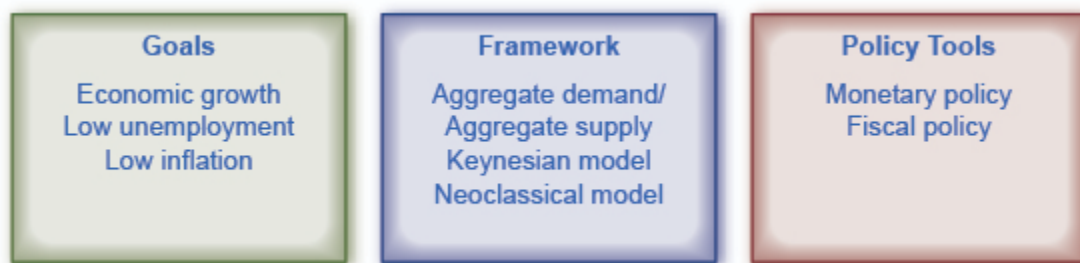
Macroeconomics focuses on the economy as a whole (or on whole economies as they interact). What causes recessions? What makes unemployment stay high when recessions are supposed to be over? Why do some countries grow faster than others? Why do some countries have higher standards of living than others? These are all questions that macroeconomics addresses. Macroeconomics involves adding up the economic activity of all households and all businesses in all markets to get the overall demand and supply in the economy. However, when we do that, something curious happens. It is not unusual that what results at the macro level is different from the sum of the microeconomic parts. Indeed, what seems sensible from a microeconomic point of view can have unexpected or counterproductive results at the macroeconomic level. Imagine that you are sitting at an event with a large audience, like a live concert or a basketball game. A few people decide that they want a better view, and so they stand up. However, when these people stand up, they block the view for other people, and the others need to stand up as well if they wish to see. Eventually, nearly everyone is standing up, and as a result, no one can see much better than before. The rational decision of some individuals at the

micro level—to stand up for a better view—ended up being self-defeating at the macro level. This is not macroeconomics, but it is an apt analogy.

Macroeconomics is a rather massive subject. How are we going to tackle it? [\[link\]](#) illustrates the structure we will use. We will study macroeconomics from three different perspectives:

1. What are the macroeconomic goals? (Macroeconomics as a discipline does not have goals, but we do have goals for the macro economy.)
2. What are the frameworks economists can use to analyze the macroeconomy?
3. Finally, what are the policy tools governments can use to manage the macroeconomy?

Macroeconomic Goals, Framework, and Policies



This chart shows what macroeconomics is about. The box on the left indicates a consensus of what are the most important goals for the macro economy, the middle box lists the frameworks economists use to analyze macroeconomic changes (such as inflation or recession), and the box on the right indicates the two tools the federal government uses to influence the macro economy.

Goals

In thinking about the overall health of the macroeconomy, it is useful to consider three primary goals: economic growth, low unemployment, and low inflation.

- Economic growth ultimately determines the prevailing standard of living in a country. Economic growth is measured by the percentage change in real (inflation-adjusted) gross domestic product. A growth rate of more than 3% is considered good.
- Unemployment, as measured by the unemployment rate, is the percentage of people in the labor force who do not have a job. When people lack jobs, the economy is wasting a precious resource—labor, and the result is lower goods and services produced. Unemployment, however, is more than a statistic—it represents people’s livelihoods. While measured unemployment is unlikely to ever be zero, a measured unemployment rate of 5% or less is considered low (good).
- Inflation is a sustained increase in the overall level of prices, and is measured by the consumer price index. If many people face a situation where the prices that they pay for food, shelter, and healthcare are rising much faster than the wages they receive for their labor, there will be widespread unhappiness as their standard of living declines. For that reason, low inflation—an inflation rate of 1–2%—is a major goal.

Frameworks

As you learn in the micro part of this book, principal tools used by economists are theories and models (see [Welcome to Economics!](#) for more on this). In microeconomics, we used the theories of supply and demand; in macroeconomics, we use the theories of aggregate demand (AD) and aggregate supply (AS). This book presents two perspectives on macroeconomics: the Neoclassical perspective and the Keynesian perspective, each of which has its own version of AD and AS. Between the two perspectives, you will obtain a good understanding of what drives the macroeconomy.

Policy Tools

National governments have two tools for influencing the macroeconomy. The first is monetary policy, which involves managing the money supply

and interest rates. The second is fiscal policy, which involves changes in government spending/purchases and taxes.

Each of the items in [\[link\]](#) will be explained in detail in one or more other chapters. As you learn these things, you will discover that the goals and the policy tools are in the news almost every day.

Measuring the Size of the Economy: Gross Domestic Product

By the end of this section, you will be able to:

- Identify the components of GDP on the demand side and on the supply side
- Evaluate how gross domestic product (GDP) is measured
- Contrast and calculate GDP, net exports, and net national product

Macroeconomics is an empirical subject, so the first step toward understanding it is to measure the economy.

How large is the U.S. economy? The size of a nation's overall economy is typically measured by its **gross domestic product (GDP)**, which is the value of all final goods and services produced within a country in a given year. The measurement of GDP involves counting up the production of millions of different goods and services—smart phones, cars, music downloads, computers, steel, bananas, college educations, and all other new goods and services produced in the current year—and summing them into a total dollar value. This task is straightforward: take the quantity of everything produced, multiply it by the price at which each product sold, and add up the total. In 2014, the U.S. GDP totaled \$17.4 trillion, the largest GDP in the world.

Each of the market transactions that enter into GDP must involve both a buyer and a seller. The GDP of an economy can be measured either by the total dollar value of what is purchased in the economy, or by the total dollar value of what is produced. There is even a third way, as we will explain later.

GDP Measured by Components of Demand

Who buys all of this production? This demand can be divided into four main parts: consumer spending (consumption), business spending (investment), government spending on goods and services, and spending on net exports. (See the following Clear It Up feature to understand what is meant by investment.) [\[link\]](#) shows how these four components added up to the GDP in 2014. [\[link\]](#) (a) shows the levels of consumption, investment, and government purchases over time, expressed as a percentage of GDP, while [\[link\]](#) (b) shows the levels of exports and imports as a percentage of GDP over time. A few patterns about each of these components are worth noticing. [\[link\]](#) shows the components of GDP from the demand side.

	Components of GDP on the Demand Side (in trillions of dollars)	Percentage of Total
Consumption	\$11.9	68.4%

	Components of GDP on the Demand Side (in trillions of dollars)	Percentage of Total
Investment	\$2.9	16.7%
Government	\$3.2	18.4%
Exports	\$2.3	13.2%
Imports	−\$2.9	−16.7%
Total GDP	\$17.4	100%

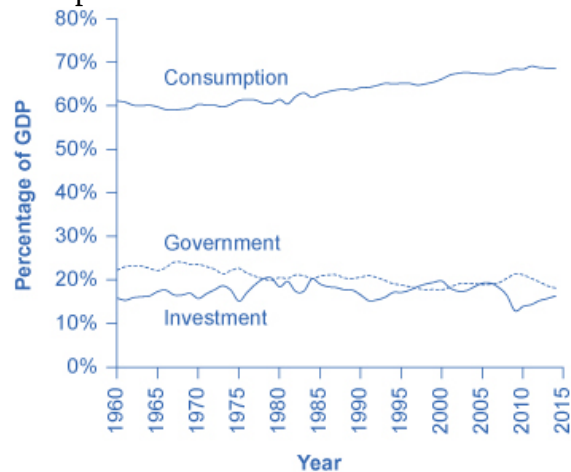
Components of U.S. GDP in 2014: From the Demand Side(Source: http://bea.gov/iTable/index_nipa.cfm)

Note:

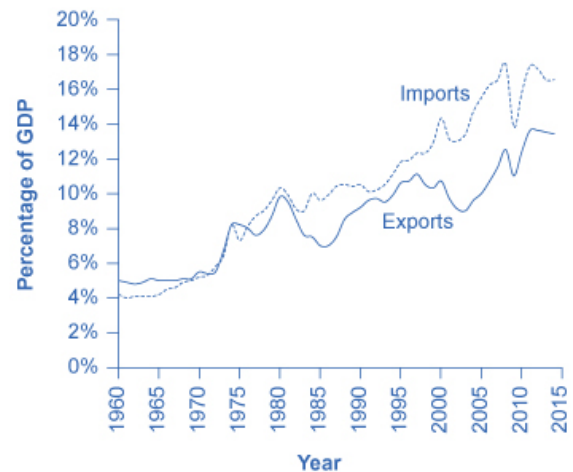
What is meant by the word “investment”?

What do economists mean by investment, or business spending? In calculating GDP, investment does not refer to the purchase of stocks and bonds or the trading of financial assets. It refers to the purchase of new capital goods, that is, new commercial real estate (such as buildings, factories, and stores) and equipment, residential housing construction, and inventories. Inventories that are produced this year are included in this year’s GDP—even if they have not yet sold. From the accountant’s perspective, it is as if the firm invested in its own inventories. Business investment in 2014 was almost \$3 trillion, according to the Bureau of Economic Analysis.

Components of GDP on the Demand Side



(a) Demand from consumption, investment, and government



(b) Imports and exports

(a) Consumption is about two-thirds of GDP, but it moves relatively little over time. Business investment hovers around 15% of GDP, but it increases and declines more than consumption. Government spending on goods and services is around 20% of GDP. (b) Exports are added to total demand for goods and services, while imports are subtracted from total demand. If exports exceed imports, as in most of the 1960s and 1970s in the U.S. economy, a trade surplus exists. If imports exceed exports, as in recent years, then a trade deficit exists. (Source: http://bea.gov/iTable/index_nipa.cfm)

Consumption expenditure by households is the largest component of GDP, accounting for about two-thirds of the GDP in any year. This tells us that consumers' spending decisions are a major driver of the economy. However, consumer spending is a gentle elephant: when viewed over time, it does not jump around too much.

Investment expenditure refers to purchases of physical plant and equipment, primarily by businesses. If Starbucks builds a new store, or Amazon buys robots, these expenditures are counted under business investment. Investment demand is far smaller than consumption demand, typically accounting for only about 15–18% of GDP, but it is very important for the economy because this is where jobs are created. However, it fluctuates more noticeably than consumption. Business investment is volatile; new technology or a new product can spur business investment, but then confidence can drop and business investment can pull back sharply.

If you have noticed any of the infrastructure projects (new bridges, highways, airports) launched during the recession of 2009, you have seen how important government spending can be for the economy. Government expenditure in the United States is about 20% of GDP, and includes spending by all three levels of government: federal, state, and local. The only part of government spending counted in demand is government purchases of goods or services produced in the economy. Examples include the government buying a new fighter jet for the Air Force (federal government spending), building a new highway (state government spending), or a new school (local government spending). A significant portion of government budgets are transfer payments, like unemployment benefits, veteran's benefits, and Social Security payments to retirees. These payments are excluded from GDP because the government does not receive a new good or service in return or exchange. Instead they are transfers of income from taxpayers to others. If you are curious about the awesome undertaking of adding up GDP, read the following Clear It Up feature.

Note:

How do statisticians measure GDP?

Government economists at the Bureau of Economic Analysis (BEA), within the U.S. Department of Commerce, piece together estimates of GDP from a variety of sources. Once every five years, in the second and seventh year of each decade, the Bureau of the Census carries out a detailed census of businesses throughout the United States. In between,

the Census Bureau carries out a monthly survey of retail sales. These figures are adjusted with foreign trade data to account for exports that are produced in the United States and sold abroad and for imports that are produced abroad and sold here. Once every ten years, the Census Bureau conducts a comprehensive survey of housing and residential finance. Together, these sources provide the main basis for figuring out what is produced for consumers.

For investment, the Census Bureau carries out a monthly survey of construction and an annual survey of expenditures on physical capital equipment.

For what is purchased by the federal government, the statisticians rely on the U.S.

Department of the Treasury. An annual Census of Governments gathers information on state and local governments. Because a lot of government spending at all levels involves hiring people to provide services, a large portion of government spending is also tracked through payroll records collected by state governments and by the Social Security Administration.

With regard to foreign trade, the Census Bureau compiles a monthly record of all import and export documents. Additional surveys cover transportation and travel, and adjustment is made for financial services that are produced in the United States for foreign customers.

Many other sources contribute to the estimates of GDP. Information on energy comes from the U.S. Department of Transportation and Department of Energy. Information on healthcare is collected by the Agency for Health Care Research and Quality. Surveys of landlords find out about rental income. The Department of Agriculture collects statistics on farming.

All of these bits and pieces of information arrive in different forms, at different time intervals. The BEA melds them together to produce estimates of GDP on a quarterly basis (every three months). These numbers are then “annualized” by multiplying by four. As more information comes in, these estimates are updated and revised. The “advance” estimate of GDP for a certain quarter is released one month after a quarter. The “preliminary” estimate comes out one month after that. The “final” estimate is published one month later, but it is not actually final. In July, roughly updated estimates for the previous calendar year are released. Then, once every five years, after the results of the latest detailed five-year business census have been processed, the BEA revises all of the past estimates of GDP according to the newest methods and data, going all the way back to 1929.

Note:

Visit this [website](#) to read FAQs on the BEA site. You can even email your own questions!



When thinking about the demand for domestically produced goods in a global economy, it is important to count spending on exports—domestically produced goods that are sold abroad. By the same token, we must also subtract spending on imports—goods produced in other countries that are purchased by residents of this country. The net export component of GDP is equal to the dollar value of exports (X) minus the dollar value of imports (M), (X – M). The gap between exports and imports is called the **trade balance**. If a country’s exports are larger than its imports, then a country is said to have a **trade surplus**. In the United States, exports typically exceeded imports in the 1960s and 1970s, as shown in [\[link\]](#) (b).

Since the early 1980s, imports have typically exceeded exports, and so the United States has experienced a **trade deficit** in most years. Indeed, the trade deficit grew quite large in the late 1990s and in the mid-2000s. [\[link\]](#) (b) also shows that imports and exports have both risen substantially in recent decades, even after the declines during the Great Recession between 2008 and 2009. As noted before, if exports and imports are equal, foreign trade has no effect on total GDP. However, even if exports and imports are balanced overall, foreign trade might still have powerful effects on particular industries and workers by causing nations to shift workers and physical capital investment toward one industry rather than another.

Based on these four components of demand, GDP can be measured as:

Equation:

$$\begin{aligned} \text{GDP} &= \text{Consumption} + \text{Investment} + \text{Government} + \text{Trade balance} \\ \text{GDP} &= C + I + G + (X - M) \end{aligned}$$

Understanding how to measure GDP is important for analyzing connections in the macro economy and for thinking about macroeconomic policy tools.

GDP Measured by What is Produced

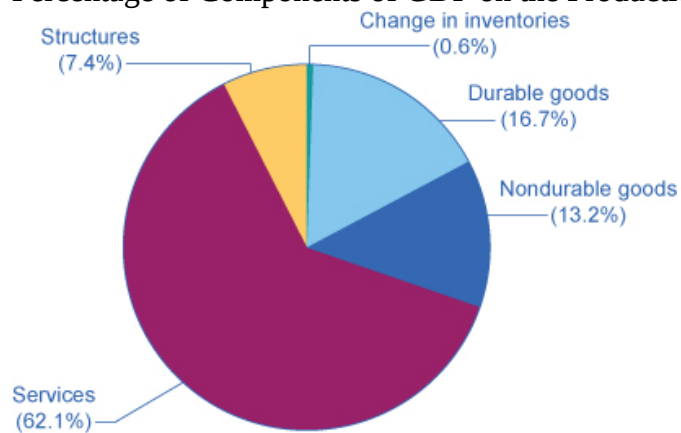
Everything that is purchased must be produced first. [\[link\]](#) breaks down what is produced into five categories: **durable goods**, **nondurable goods**, **services**, **structures**, and the change in **inventories**. Before going into detail about these categories, notice that total GDP measured according to what is produced is exactly the same as the GDP measured by looking at the five components of demand. [\[link\]](#) provides a visual representation of this information.

	Components of GDP on the Supply Side (in trillions of dollars)	Percentage of Total
Goods		

	Components of GDP on the Supply Side (in trillions of dollars)	Percentage of Total
Durable goods	\$2.9	16.7%
Nondurable goods	\$2.3	13.2%
Services	\$10.8	62.1%
Structures	\$1.3	7.4%
Change in inventories	\$0.1	0.6%
Total GDP	\$17.4	100%

Components of U.S. GDP on the Production Side, 2014(Source: http://bea.gov/iTable/index_nipa.cfm)

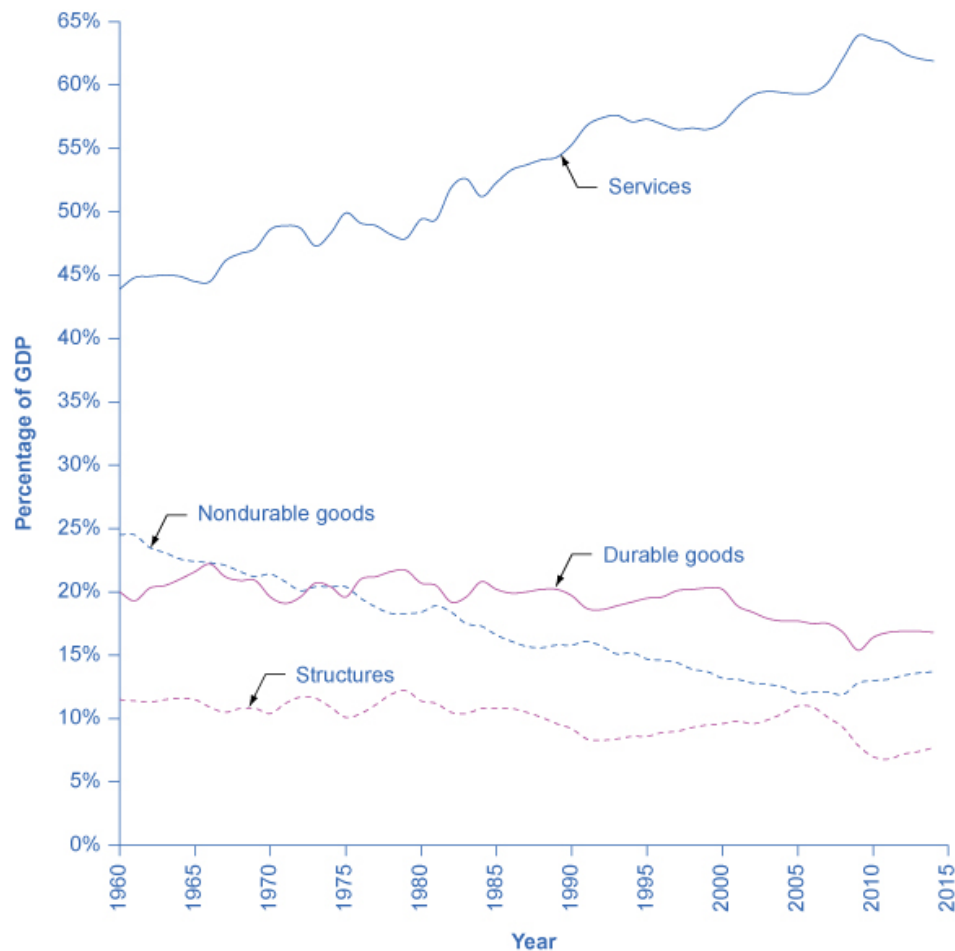
Percentage of Components of GDP on the Production Side



Services make up over half of the production side components of GDP in the United States.

Since every market transaction must have both a buyer and a seller, GDP must be the same whether measured by what is demanded or by what is produced. [\[link\]](#) shows these components of what is produced, expressed as a percentage of GDP, since 1960.

Types of Production



Services are the largest single component of total supply, representing over half of GDP. Nondurable goods used to be larger than durable goods, but in recent years, nondurable goods have been dropping closer to durable goods, which is about 20% of GDP. Structures hover around 10% of GDP. The change in inventories, the final component of aggregate supply, is not shown here; it is typically less than 1% of GDP.

In thinking about what is produced in the economy, many non-economists immediately focus on solid, long-lasting goods, like cars and computers. By far the largest part of GDP, however, is services. Moreover, services have been a growing share of GDP over time. A detailed breakdown of the leading service industries would include healthcare, education, and legal and financial services. It has been decades since most of the U.S. economy involved making solid objects. Instead, the most common jobs in a modern economy involve a worker looking at pieces of paper or a computer screen; meeting with co-workers, customers, or suppliers; or making phone calls.

Even within the overall category of goods, long-lasting durable goods like cars and refrigerators are about the same share of the economy as short-lived nondurable goods like food and clothing. The category of structures includes everything from homes, to office buildings, shopping malls, and factories. Inventories is a small category that refers to the goods that have been produced by one business but have not yet been sold to consumers, and are still sitting in warehouses and on shelves. The amount of inventories sitting on shelves tends to decline if business is better than expected, or to rise if business is worse than expected.

The Problem of Double Counting

GDP is defined as the current value of all final goods and services produced in a nation in a year. What are final goods? They are goods at the furthest stage of production at the end of a year. Statisticians who calculate GDP must avoid the mistake of **double counting**, in which output is counted more than once as it travels through the stages of production. For example, imagine what would happen if government statisticians first counted the value of tires produced by a tire manufacturer, and then counted the value of a new truck sold by an automaker that contains those tires. In this example, the value of the tires would have been counted twice—because the price of the truck includes the value of the tires.

To avoid this problem, which would overstate the size of the economy considerably, government statisticians count just the value of **final goods and services** in the chain of production that are sold for consumption, investment, government, and trade purposes. **Intermediate goods**, which are goods that go into the production of other goods, are excluded from GDP calculations. From the example above, only the value of the Ford truck will be counted. The value of what businesses provide to other businesses is captured in the final products at the end of the production chain.

The concept of GDP is fairly straightforward: it is just the dollar value of all final goods and services produced in the economy in a year. In our decentralized, market-oriented economy, actually calculating the more than \$16 trillion-dollar U.S. GDP—along with how it is changing every few months—is a full-time job for a brigade of government statisticians.

What is Counted in GDP	What is not included in GDP
Consumption	Intermediate goods
Business investment	Transfer payments and non-market activities

What is Counted in GDP	What is not included in GDP
Government spending on goods and services	Used goods
Net exports	Illegal goods

Counting GDP

Notice the items that are not counted into GDP, as outlined in [\[link\]](#). The sales of used goods are not included because they were produced in a previous year and are part of that year's GDP. The entire underground economy of services paid “under the table” and illegal sales should be counted, but is not, because it is impossible to track these sales. In a recent study by Friedrich Schneider of shadow economies, the underground economy in the United States was estimated to be 6.6% of GDP, or close to \$2 trillion dollars in 2013 alone. Transfer payments, such as payment by the government to individuals, are not included, because they do not represent production. Also, production of some goods—such as home production as when you make your breakfast—is not counted because these goods are not sold in the marketplace.

Note:

Visit this [website](#) to read about the “New Underground Economy.”



Other Ways to Measure the Economy

Besides GDP, there are several different but closely related ways of measuring the size of the economy. We mentioned above that GDP can be thought of as total production and as total purchases. It can also be thought of as total income since anything produced and sold produces income.

One of the closest cousins of GDP is the **gross national product (GNP)**. GDP includes only what is produced within a country's borders. GNP adds what is produced by domestic businesses and labor abroad, and subtracts out any payments sent home to other countries by foreign labor and businesses located in the United States. In other words, GNP is based more on the production of citizens and firms of a country, wherever they are located, and GDP is

based on what happens within the geographic boundaries of a certain country. For the United States, the gap between GDP and GNP is relatively small; in recent years, only about 0.2%. For small nations, which may have a substantial share of their population working abroad and sending money back home, the difference can be substantial.

Net national product (NNP) is calculated by taking GNP and then subtracting the value of how much physical capital is worn out, or reduced in value because of aging, over the course of a year. The process by which capital ages and loses value is called **depreciation**. The NNP can be further subdivided into **national income**, which includes all income to businesses and individuals, and personal income, which includes only income to people.

For practical purposes, it is not vital to memorize these definitions. However, it is important to be aware that these differences exist and to know what statistic you are looking at, so that you do not accidentally compare, say, GDP in one year or for one country with GNP or NNP in another year or another country. To get an idea of how these calculations work, follow the steps in the following Work It Out feature.

Note:

Calculating GDP, Net Exports, and NNP

Based on the information in [\[link\]](#):

- a. What is the value of GDP?
- b. What is the value of net exports?
- c. What is the value of NNP?

Government purchases	\$120 billion
Depreciation	\$40 billion
Consumption	\$400 billion
Business Investment	\$60 billion
Exports	\$100 billion
Imports	\$120 billion
Income receipts from rest of the world	\$10 billion
Income payments to rest of the world	\$8 billion

Step 1. To calculate GDP use the following formula:

Equation:

$$\begin{aligned}\text{GDP} &= \text{Consumption} + \text{Investment} + \text{Government spending} + (\text{Exports} - \text{Imports}) \\ &= C + I + G + (X - M) \\ &= \$400 + \$60 + \$120 + (\$100 - \$120) \\ &= \$560 \text{ billion}\end{aligned}$$

Step 2. To calculate net exports, subtract imports from exports.

Equation:

$$\begin{aligned}\text{Net exports} &= X - M \\ &= \$100 - \$120 \\ &= -\$20 \text{ billion}\end{aligned}$$

Step 3. To calculate NNP, use the following formula:

Equation:

$$\begin{aligned}\text{NNP} &= \text{GDP} + \text{Income receipts from the rest of the world} \\ &\quad - \text{Income payments to the rest of the world} - \text{Depreciation} \\ &= \$560 + \$10 - \$8 - \$40 \\ &= \$522 \text{ billion}\end{aligned}$$

Key Concepts and Summary

The size of a nation's economy is commonly expressed as its gross domestic product (GDP), which measures the value of the output of all goods and services produced within the country in a year. GDP is measured by taking the quantities of all goods and services produced, multiplying them by their prices, and summing the total. Since GDP measures what is bought and sold in the economy, it can be measured either by the sum of what is purchased in the economy or what is produced.

Demand can be divided into consumption, investment, government, exports, and imports. What is produced in the economy can be divided into durable goods, nondurable goods, services, structures, and inventories. To avoid double counting, GDP counts only final output of goods and services, not the production of intermediate goods or the value of labor in the chain of production.

Self-Check Questions

Exercise:**Problem:**

Country A has export sales of \$20 billion, government purchases of \$1,000 billion, business investment is \$50 billion, imports are \$40 billion, and consumption spending is \$2,000 billion. What is the dollar value of GDP?

Solution:

GDP is $C + I + G + (X - M)$. $GDP = \$2,000 \text{ billion} + \$50 \text{ billion} + \$1,000 \text{ billion} + (\$20 \text{ billion} - \$40 \text{ billion}) = \$3,030$

Exercise:

Problem: Which of the following are included in GDP, and which are not?

- a. The cost of hospital stays
 - b. The rise in life expectancy over time
 - c. Child care provided by a licensed day care center
 - d. Child care provided by a grandmother
 - e. The sale of a used car
 - f. The sale of a new car
 - g. The greater variety of cheese available in supermarkets
 - h. The iron that goes into the steel that goes into a refrigerator bought by a consumer.
-

Solution:

- a. Hospital stays are part of GDP.
- b. Changes in life expectancy are not market transactions and not part of GDP.
- c. Child care that is paid for is part of GDP.
- d. If Grandma gets paid and reports this as income, it is part of GDP, otherwise not.
- e. A used car is not produced this year, so it is not part of GDP.
- f. A new car is part of GDP.
- g. Variety does not count in GDP, where the cheese could all be cheddar.
- h. The iron is not counted because it is an intermediate good.

Review Questions**Exercise:**

Problem: What are the main components of measuring GDP with what is demanded?

Exercise:

Problem: What are the main components of measuring GDP with what is produced?

Exercise:

Problem:

Would you usually expect GDP as measured by what is demanded to be greater than GDP measured by what is supplied, or the reverse?

Exercise:

Problem: Why must double counting be avoided when measuring GDP?

Critical Thinking Question

Exercise:

Problem:

U.S. macroeconomic data are thought to be among the best in the world. Given what you learned in the [Clear It Up](#) "How do statisticians measure GDP?", does this surprise you? Or does this simply reflect the complexity of a modern economy?

Exercise:

Problem: What does GDP not tell us about the economy?

Problem

Exercise:

Problem:

Last year, a small nation with abundant forests cut down \$200 worth of trees. \$100 worth of trees were then turned into \$150 worth of lumber. \$100 worth of that lumber was used to produce \$250 worth of bookshelves. Assuming the country produces no other outputs, and there are no other inputs used in the production of trees, lumber, and bookshelves, what is this nation's GDP? In other words, what is the value of the final goods produced including trees, lumber and bookshelves?

References

U.S. Department of Commerce: Bureau of Economic Analysis. "National data: National Income and Product Accounts Tables." <http://bea.gov/iTable/iTable.cfm?ReqID=9&step=1>.

U.S. Department of Commerce: United States Census Bureau. "Census of Governments: 2012 Census of Governments." <http://www.census.gov/govs/cog/>.

United States Department of Transportation. "About DOT." Last modified March 2, 2012. <http://www.dot.gov/about>.

U.S. Department of Energy. "Energy.gov." <http://energy.gov/>.

U.S. Department of Health & Human Services. "Agency for Healthcare Research and Quality." <http://www.ahrq.gov/>.

United States Department of Agriculture. "USDA." <http://www.usda.gov/wps/portal/usda/usdahome>.

Schneider, Friedrich. Department of Economics. "Size and Development of the Shadow Economy of 31 European and 5 other OECD Countries from 2003 to 2013: A Further Decline." *Johannes Kepler University*. Last modified April 5, 2013. http://www.econ.jku.at/members/Schneider/files/publications/2013/ShadEcEurope31_Jan2013.pdf.

Glossary

depreciation

the process by which capital ages over time and therefore loses its value

double counting

a potential mistake to be avoided in measuring GDP, in which output is counted more than once as it travels through the stages of production

durable good

long-lasting good like a car or a refrigerator

final good and service

output used directly for consumption, investment, government, and trade purposes; contrast with "intermediate good"

gross domestic product (GDP)

the value of the output of all goods and services produced within a country in a year

gross national product (GNP)

includes what is produced domestically and what is produced by domestic labor and business abroad in a year

intermediate good

output provided to other businesses at an intermediate stage of production, not for final users; contrast with "final good and service"

inventory

good that has been produced, but not yet been sold

national income

includes all income earned: wages, profits, rent, and profit income

net national product (NNP)

GDP minus depreciation

nondurable good

short-lived good like food and clothing

service

product which is intangible (in contrast to goods) such as entertainment, healthcare, or education

structure

building used as residence, factory, office building, retail store, or for other purposes

trade balance

gap between exports and imports

trade deficit

exists when a nation's imports exceed its exports and is calculated as $\text{imports} - \text{exports}$

trade surplus

exists when a nation's exports exceed its imports and is calculated as $\text{exports} - \text{imports}$

Adjusting Nominal Values to Real Values

By the end of this section, you will be able to:

- Contrast nominal GDP and real GDP
- Explain GDP deflator
- Calculate real GDP based on nominal GDP values

When examining economic statistics, there is a crucial distinction worth emphasizing. The distinction is between nominal and real measurements, which refer to whether or not inflation has distorted a given statistic. Looking at economic statistics without considering inflation is like looking through a pair of binoculars and trying to guess how close something is: unless you know how strong the lenses are, you cannot guess the distance very accurately. Similarly, if you do not know the rate of inflation, it is difficult to figure out if a rise in GDP is due mainly to a rise in the overall level of prices or to a rise in quantities of goods produced. The **nominal value** of any economic statistic means the statistic is measured in terms of actual prices that exist at the time. The **real value** refers to the same statistic after it has been adjusted for inflation. Generally, it is the real value that is more important.

Converting Nominal to Real GDP

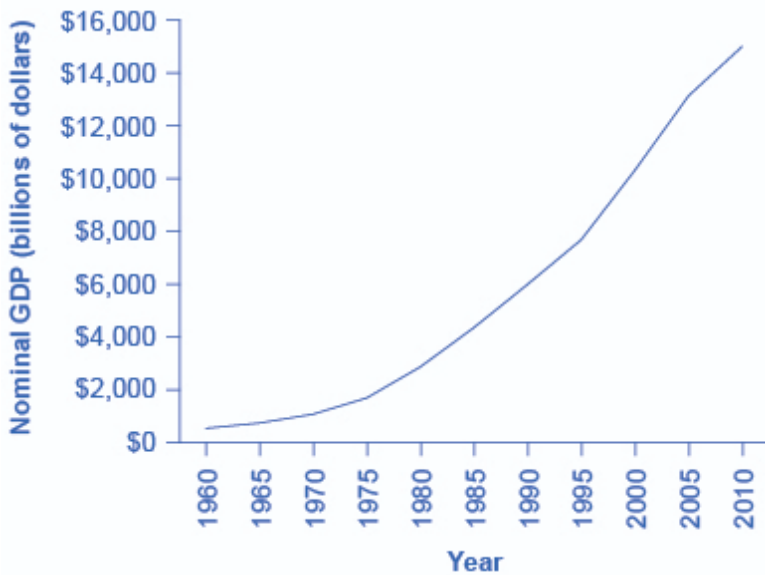
[\[link\]](#) shows U.S. GDP at five-year intervals since 1960 in nominal dollars; that is, GDP measured using the actual market prices prevailing in each stated year. This data is also reflected in the graph shown in [\[link\]](#)

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)
1960	543.3	19.0

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)
1965	743.7	20.3
1970	1,075.9	24.8
1975	1,688.9	34.1
1980	2,862.5	48.3
1985	4,346.7	62.3
1990	5,979.6	72.7
1995	7,664.0	81.7
2000	10,289.7	89.0
2005	13,095.4	100.0
2010	14,958.3	110.0

U.S. Nominal GDP and the GDP Deflator(Source: www.bea.gov)

U.S. Nominal GDP, 1960–2010

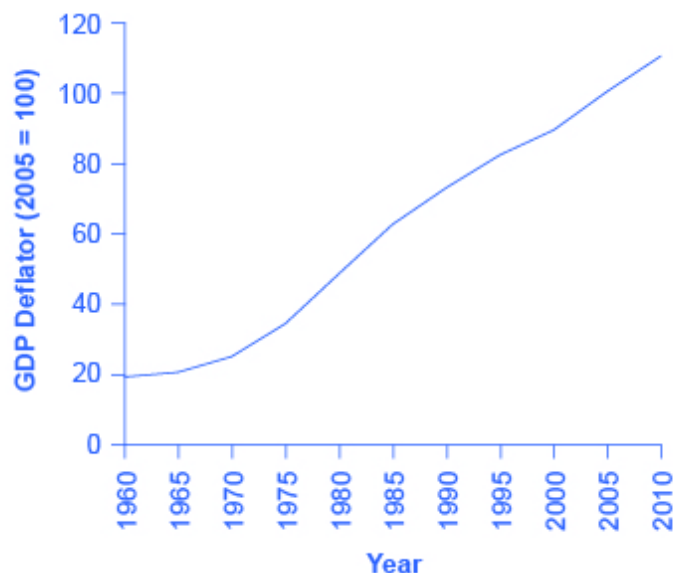


Nominal GDP values have risen exponentially from 1960 through 2010, according to the BEA.

If an unwary analyst compared nominal GDP in 1960 to nominal GDP in 2010, it might appear that national output had risen by a factor of twenty-seven over this time (that is, GDP of \$14,958 billion in 2010 divided by GDP of \$543 billion in 1960). This conclusion would be highly misleading. Recall that nominal GDP is defined as the quantity of every good or service produced multiplied by the price at which it was sold, summed up for all goods and services. In order to see how much production has actually increased, we need to extract the effects of higher prices on nominal GDP. This can be easily done, using the GDP deflator.

GDP deflator is a price index measuring the average prices of all goods and services included in the economy. We explore price indices in detail and how they are computed in [Inflation](#), but this definition will do in the context of this chapter. The data for the GDP deflator are given in [\[link\]](#) and shown graphically in [\[link\]](#).

U.S. GDP Deflator, 1960–2010



Much like nominal GDP, the GDP deflator has risen exponentially from 1960 through 2010. (Source: BEA)

[\[link\]](#) shows that the price level has risen dramatically since 1960. The price level in 2010 was almost six times higher than in 1960 (the deflator for 2010 was 110 versus a level of 19 in 1960). Clearly, much of the apparent growth in nominal GDP was due to inflation, not an actual change in the quantity of goods and services produced, in other words, not in real GDP. Recall that nominal GDP can rise for two reasons: an increase in output, and/or an increase in prices. What is needed is to extract the increase in prices from nominal GDP so as to measure only changes in output. After all, the dollars used to measure nominal GDP in 1960 are worth more than the inflated dollars of 1990—and the price index tells exactly how much more. This adjustment is easy to do if you understand that nominal measurements are in value terms, where

Equation:

$$\text{Value} = \text{Price} \times \text{Quantity}$$

or

$$\text{Nominal GDP} = \text{GDP Deflator} \times \text{Real GDP}$$

Let's look at an example at the micro level. Suppose the t-shirt company, Coolshirts, sells 10 t-shirts at a price of \$9 each.

Equation:

$$\begin{aligned}\text{Coolshirt's nominal revenue from sales} &= \text{Price} \times \text{Quantity} \\ &= \$9 \times 10 \\ &= \$90\end{aligned}$$

Then,

Equation:

$$\begin{aligned}\text{Coolshirt's real income} &= \frac{\text{Nominal revenue}}{\text{Price}} \\ &= \frac{\$90}{\$9} \\ &= 10\end{aligned}$$

In other words, when we compute “real” measurements we are trying to get at actual quantities, in this case, 10 t-shirts.

With GDP, it is just a tiny bit more complicated. We start with the same formula as above:

Equation:

$$\text{Real GDP} = \frac{\text{Nominal GDP}}{\text{Price Index}}$$

For reasons that will be explained in more detail below, mathematically, a price index is a two-digit decimal number like 1.00 or 0.85 or 1.25. Because some people have trouble working with decimals, when the price index is published, it has traditionally been multiplied by 100 to get integer numbers like 100, 85, or 125. What this means is that when we “deflate” nominal figures to get real figures (by dividing the nominal by the price index). We also need to remember to divide the published price index by 100 to make the math work. So the formula becomes:

Equation:

$$\text{Real GDP} = \frac{\text{Nominal GDP}}{\text{Price Index} / 100}$$

Now read the following Work It Out feature for more practice calculating real GDP.

Note:**Computing GDP**

It is possible to use the data in [\[link\]](#) to compute real GDP.

Step 1. Look at [\[link\]](#), to see that, in 1960, nominal GDP was \$543.3 billion and the price index (GDP deflator) was 19.0.

Step 2. To calculate the real GDP in 1960, use the formula:

Equation:

$$\begin{aligned}\text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$543.3 \text{ billion}}{19 / 100} \\ &= \$2,859.5 \text{ billion}\end{aligned}$$

We'll do this in two parts to make it clear. First adjust the price index: 19 divided by 100 = 0.19. Then divide into nominal GDP: \$543.3 billion / 0.19 = \$2,859.5 billion.

Step 3. Use the same formula to calculate the real GDP in 1965.

Equation:

$$\begin{aligned}\text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$743.7 \text{ billion}}{20.3 / 100} \\ &= \$3,663.5 \text{ billion}\end{aligned}$$

Step 4. Continue using this formula to calculate all of the real GDP values from 1960 through 2010. The calculations and the results are shown in [\[link\]](#).

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)	Calculations	Real GDP (billions of 2005 dollars)
1960	543.3	19.0	$543.3 / (19.0/100)$	2859.5
1965	743.7	20.3	$743.7 / (20.3/100)$	3663.5
1970	1075.9	24.8	$1,075.9 / (24.8/100)$	4338.3
1975	1688.9	34.1	$1,688.9 / (34.1/100)$	4952.8
1980	2862.5	48.3	$2,862.5 / (48.3/100)$	5926.5
1985	4346.7	62.3	$4,346.7 / (62.3/100)$	6977.0
1990	5979.6	72.7	$5,979.6 / (72.7/100)$	8225.0
1995	7664.0	82.0	$7,664 / (82.0/100)$	9346.3
2000	10289.7	89.0	$10,289.7 / (89.0/100)$	11561.5
2005	13095.4	100.0	$13,095.4 / (100.0/100)$	13095.4

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)	Calculations	Real GDP (billions of 2005 dollars)
2010	14958.3	110.0	14,958.3 / (110.0/100)	13598.5

Converting Nominal to Real GDP(Source: Bureau of Economic Analysis, www.bea.gov)

There are a couple things to notice here. Whenever you compute a real statistic, one year (or period) plays a special role. It is called the base year (or base period). The base year is the year whose prices are used to compute the real statistic. When we calculate real GDP, for example, we take the quantities of goods and services produced in each year (for example, 1960 or 1973) and multiply them by their prices in the base year (in this case, 2005), so we get a measure of GDP that uses prices that do not change from year to year. That is why real GDP is labeled “Constant Dollars” or “2005 Dollars,” which means that real GDP is constructed using prices that existed in 2005. The formula used is:

Equation:

$$\text{GDP deflator} = \frac{\text{Nominal GDP}}{\text{Real GDP}} \times 100$$

Rearranging the formula and using the data from 2005:

Equation:

$$\begin{aligned} \text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$13,095.4 \text{ billion}}{100 / 100} \\ &= \$13,095.4 \text{ billion} \end{aligned}$$

Comparing real GDP and nominal GDP for 2005, you see they are the same. This is no accident. It is because 2005 has been chosen as the “base year” in this example. Since the price index in the base year always has a

value of 100 (by definition), nominal and real GDP are always the same in the base year.

Look at the data for 2010.

Equation:

$$\begin{aligned}\text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$14,958.3 \text{ billion}}{110 / 100} \\ &= \$13,598.5 \text{ billion}\end{aligned}$$

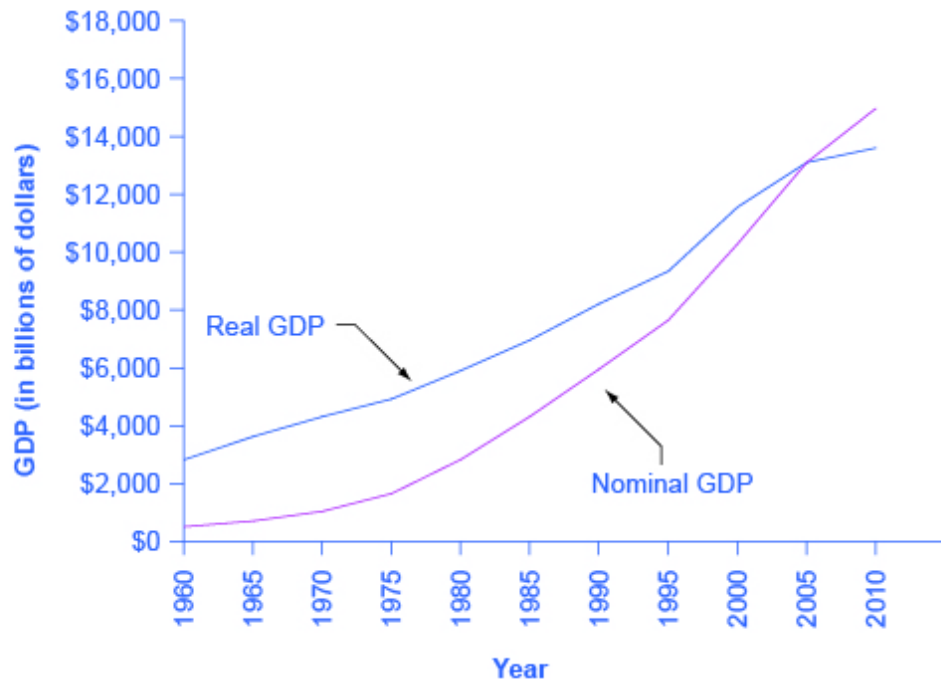
Use this data to make another observation: As long as inflation is positive, meaning prices increase on average from year to year, real GDP should be less than nominal GDP in any year after the base year. The reason for this should be clear: The value of nominal GDP is “inflated” by inflation.

Similarly, as long as inflation is positive, real GDP should be greater than nominal GDP in any year before the base year.

[\[link\]](#) shows the U.S. nominal and real GDP since 1960. Because 2005 is the base year, the nominal and real values are exactly the same in that year.

However, over time, the rise in nominal GDP looks much larger than the rise in real GDP (that is, the nominal GDP line rises more steeply than the real GDP line), because the rise in nominal GDP is exaggerated by the presence of inflation, especially in the 1970s.

U.S. Nominal and Real GDP, 1960–2012



The red line measures U.S. GDP in nominal dollars. The black line measures U.S. GDP in real dollars, where all dollar values have been converted to 2005 dollars. Since real GDP is expressed in 2005 dollars, the two lines cross in 2005. However, real GDP will appear higher than nominal GDP in the years before 2005, because dollars were worth less in 2005 than in previous years. Conversely, real GDP will appear lower in the years after 2005, because dollars were worth more in 2005 than in later years.

Let's return to the question posed originally: How much did GDP increase in real terms? What was the rate of growth of real GDP from 1960 to 2010? To find the real growth rate, we apply the formula for percentage change:

Equation:

$$\frac{2010 \text{ real GDP} - 1960 \text{ real GDP}}{1960 \text{ real GDP}} \times 100 = \% \text{ change}$$

$$\frac{13,598.5 - 2,859.5}{2,859.5} \times 100 = 376\%$$

In other words, the U.S. economy has increased real production of goods and services by nearly a factor of four since 1960. Of course, that understates the material improvement since it fails to capture improvements in the quality of products and the invention of new products.

There is a quicker way to answer this question approximately, using another math trick. Because:

Equation:

$$\text{Nominal} = \text{Price} \times \text{Quantity}$$

$$\% \text{ change in Nominal} = \% \text{ change in Price} + \% \text{ change in Quantity}$$

OR

$$\% \text{ change in Quantity} = \% \text{ change in Nominal} - \% \text{ change in Price}$$

Therefore, the growth rate of real GDP (% change in quantity) equals the growth rate in nominal GDP (% change in value) minus the inflation rate (% change in price).

Note that using this equation provides an approximation for small changes in the levels. For more accurate measures, one should use the first formula shown.

Key Concepts and Summary

The nominal value of an economic statistic is the commonly announced value. The real value is the value after adjusting for changes in inflation. To convert nominal economic data from several different years into real, inflation-adjusted data, the starting point is to choose a base year arbitrarily and then use a price index to convert the measurements so that they are measured in the money prevailing in the base year.

Self-Check Question

Exercise:

Problem:

Using data from [\[link\]](#) how much of the nominal GDP growth from 1980 to 1990 was real GDP and how much was inflation?

Solution:

From 1980 to 1990, real GDP grew by $(8,225.0 - 5,926.5) / (5,926.5) = 39\%$. Over the same period, prices increased by $(72.7 - 48.3) / (48.3/100) = 51\%$. So about 57% of the growth $51 / (51 + 39)$ was inflation, and the remainder: $39 / (51 + 39) = 43\%$ was growth in real GDP.

Review Questions**Exercise:****Problem:**

What is the difference between a series of economic data over time measured in nominal terms versus the same data series over time measured in real terms?

Exercise:**Problem:**

How do you convert a series of nominal economic data over time to real terms?

Critical Thinking Question**Exercise:**

Problem:

Should people typically pay more attention to their real income or their nominal income? If you choose the latter, why would that make sense in today's world? Would your answer be the same for the 1970s?

Problems**Exercise:****Problem:**

The “prime” interest rate is the rate that banks charge their best customers. Based on the nominal interest rates and inflation rates given in [\[link\]](#), in which of the years given would it have been best to be a lender? Based on the nominal interest rates and inflation rates given in [\[link\]](#), in which of the years given would it have been best to be a borrower?

Year	Prime Interest Rate	Inflation Rate
1970	7.9%	5.7%
1974	10.8%	11.0%
1978	9.1%	7.6%
1981	18.9%	10.3%

Exercise:

Problem:

A mortgage loan is a loan that a person makes to purchase a house. [\[link\]](#) provides a list of the mortgage interest rate being charged for several different years and the rate of inflation for each of those years. In which years would it have been better to be a person borrowing money from a bank to buy a home? In which years would it have been better to be a bank lending money?

Year	Mortgage Interest Rate	Inflation Rate
1984	12.4%	4.3%
1990	10%	5.4%
2001	7.0%	2.8%

Glossary

nominal value

the economic statistic actually announced at that time, not adjusted for inflation; contrast with real value

real value

an economic statistic after it has been adjusted for inflation; contrast with nominal value

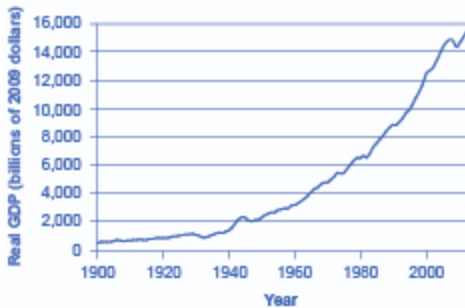
Tracking Real GDP over Time

By the end of this section, you will be able to:

- Explain recessions, depressions, peaks, and troughs
- Evaluate the importance of tracking real GDP over time

When news reports indicate that “the economy grew 1.2% in the first quarter,” the reports are referring to the percentage change in real GDP. By convention, GDP growth is reported at an annualized rate: Whatever the calculated growth in real GDP was for the quarter, it is multiplied by four when it is reported as if the economy were growing at that rate for a full year.

U.S. GDP, 1900–2014



Real GDP in the United States in 2014 was about \$16 trillion. After adjusting to remove the effects of inflation, this represents a roughly 20-fold increase in the economy’s production of goods and services since the start of the twentieth century.
(Source: bea.gov)

[\[link\]](#) shows the pattern of U.S. real GDP since 1900. The generally upward long-term path of GDP has been regularly interrupted by short-term declines. A significant decline in real GDP is called a **recession**. An especially lengthy and deep recession is called a **depression**. The severe drop in GDP that occurred during the Great Depression of the 1930s is clearly visible in the figure, as is the Great Recession of 2008–2009.

Real GDP is important because it is highly correlated with other measures of economic activity, like employment and unemployment. When real GDP rises, so does employment.

The most significant human problem associated with recessions (and their larger, uglier cousins, depressions) is that a slowdown in production means that firms need to lay off or fire some of the workers they have. Losing a job imposes painful financial and personal costs on workers, and often on their extended families as well. In addition, even those who keep their jobs are likely to find that wage raises are scanty at best—or they may even be asked to take pay cuts.

[\[link\]](#) lists the pattern of recessions and expansions in the U.S. economy since 1900. The highest point of the economy, before the recession begins, is called the **peak**; conversely, the lowest point of a recession, before a recovery begins, is called the **trough**. Thus, a recession lasts from peak to trough, and an economic upswing runs from trough to peak. The movement of the economy from peak to trough and trough to peak is called the **business cycle**. It is intriguing to notice that the three longest trough-to-peak expansions of the twentieth century have happened since 1960. The most recent recession started in December 2007 and ended formally in June 2009. This was the most severe recession since the Great Depression of the 1930's.

Trough	Peak	Months of Contraction	Months of Expansion
December 1900	September 1902	18	21
August 1904	May 1907	23	33
June 1908	January 1910	13	19
January 1912	January 1913	24	12
December 1914	August 1918	23	44
March 1919	January 1920	7	10
July 1921	May 1923	18	22
July 1924	October 1926	14	27
November 1927	August 1929	23	21
March 1933	May 1937	43	50
June 1938	February 1945	13	80

Trough	Peak	Months of Contraction	Months of Expansion
October 1945	November 1948	8	37
October 1949	July 1953	11	45
May 1954	August 1957	10	39
April 1958	April 1960	8	24
February 1961	December 1969	10	106
November 1970	November 1973	11	36
March 1975	January 1980	16	58
July 1980	July 1981	6	12
November 1982	July 1990	16	92
March 2001	November 2001	8	120
December 2007	June 2009	18	73

U.S. Business Cycles since 1900(Source:
<http://www.nber.org/cycles/main.html>)

A private think tank, the National Bureau of Economic Research (NBER), is the official tracker of business cycles for the U.S. economy. However, the effects of a severe recession often linger on after the official ending date assigned by the NBER.

Key Concepts and Summary

Over the long term, U.S. real GDP have increased dramatically. At the same time, GDP has not increased the same amount each year. The speeding up and slowing down of GDP growth represents the business cycle. When GDP declines significantly, a recession occurs. A longer and deeper decline is a depression. Recessions begin at the peak of the business cycle and end at the trough.

Self-Check Questions

Exercise:

Problem:

Without looking at [\[link\]](#), return to [\[link\]](#). If a recession is defined as a significant decline in national output, can you identify any post-1960 recessions in addition to the recession of 2008–2009? (This requires a judgment call.)

Solution:

Two other major recessions are visible in the figure as slight dips: those of 1973–1975, and 1981–1982. Two other recessions appear in the figure as a flattening of the path of real GDP. These were in 1990–1991 and 2001.

Exercise:

Problem:

According to [\[link\]](#), how often have recessions occurred since the end of World War II (1945)?

Solution:

11 recessions in approximately 70 years averages about one recession every six years.

Exercise:**Problem:**

According to [\[link\]](#), how long has the average recession lasted since the end of World War II?

Solution:

The table lists the “Months of Contraction” for each recession. Averaging these figures for the post-WWII recessions gives an average duration of 11 months, or slightly less than a year.

Exercise:**Problem:**

According to [\[link\]](#), how long has the average expansion lasted since the end of World War II?

Solution:

The table lists the “Months of Expansion.” Averaging these figures for the post-WWII expansions gives an average expansion of 60.5 months, or more than five years.

Review Question**Exercise:****Problem:**

What are the typical patterns of GDP for a high-income economy like the United States in the long run and the short run?

Critical Thinking Questions

Exercise:

Problem:

Why do you suppose that U.S. GDP is so much higher today than 50 or 100 years ago?

Exercise:

Problem:

Why do you think that GDP does not grow at a steady rate, but rather speeds up and slows down?

References

The National Bureau of Economic Research. "Information on Recessions and Recoveries, the NBER Business Cycle Dating Committee, and related topics." <http://www.nber.org/cycles/main.html>.

Glossary

business cycle

the relatively short-term movement of the economy in and out of recession

depression

an especially lengthy and deep decline in output

peak

during the business cycle, the highest point of output before a recession begins

recession

a significant decline in national output

trough

during the business cycle, the lowest point of output in a recession,
before a recovery begins

Comparing GDP among Countries

By the end of this section, you will be able to:

- Explain how GDP can be used to compare the economic welfare of different nations
- Calculate the conversion of GDP to a common currency by using exchange rates
- Calculate GDP per capita using population data

It is common to use GDP as a measure of economic welfare or standard of living in a nation. When comparing the GDP of different nations for this purpose, two issues immediately arise. First, the GDP of a country is measured in its own currency: the United States uses the U.S. dollar; Canada, the Canadian dollar; most countries of Western Europe, the euro; Japan, the yen; Mexico, the peso; and so on. Thus, comparing GDP between two countries requires converting to a common currency. A second issue is that countries have very different numbers of people. For instance, the United States has a much larger economy than Mexico or Canada, but it also has roughly three times as many people as Mexico and nine times as many people as Canada. So, if we are trying to compare standards of living across countries, we need to divide GDP by population.

Converting Currencies with Exchange Rates

To compare the GDP of countries with different currencies, it is necessary to convert to a “common denominator” using an **exchange rate**, which is the value of one currency in terms of another currency. Exchange rates are expressed either as the units of country A’s currency that need to be traded for a single unit of country B’s currency (for example, Japanese yen per British pound), or as the inverse (for example, British pounds per Japanese yen). Two types of exchange rates can be used for this purpose, market exchange rates and purchasing power parity (PPP) equivalent exchange rates. Market exchange rates vary on a day-to-day basis depending on supply and demand in foreign exchange markets. PPP-equivalent exchange rates provide a longer run measure of the exchange rate. For this reason, PPP-equivalent exchange rates are typically used for cross country comparisons of GDP. Exchange rates will be discussed in more detail in [Exchange Rates and International Capital Flows](#). The following Work It Out feature explains how to convert GDP to a common currency.

Note:**Converting GDP to a Common Currency**

Using the exchange rate to convert GDP from one currency to another is straightforward. Say that the task is to compare Brazil's GDP in 2013 of 4.8 trillion reais with the U.S. GDP of \$16.6 trillion for the same year.

Step 1. Determine the exchange rate for the specified year. In 2013, the exchange rate was 2.230 reais = \$1. (These numbers are realistic, but rounded off to simplify the calculations.)

Step 2. Convert Brazil's GDP into U.S. dollars:

Equation:

$$\begin{aligned}
 \text{Brazil's GDP in \$ U.S.} &= \frac{\text{Brazil's GDP in reais}}{\text{Exchange rate (reais/\$ U.S.)}} \\
 &= \frac{4.8 \text{ trillion reais}}{2.230 \text{ reais per \$ U.S.}} \\
 &= \$2.2 \text{ trillion}
 \end{aligned}$$

Step 3. Compare this value to the GDP in the United States in the same year. The U.S. GDP was \$16.6 trillion in 2013, which is nearly eight times that of GDP in Brazil in 2012.

Step 4. View [\[link\]](#) which shows the size of and variety of GDPs of different countries in 2013, all expressed in U.S. dollars. Each is calculated using the process explained above.

Country	GDP in Billions of Domestic Currency		Domestic Currency/U.S. Dollars (PPP Equivalent)	GDP (in billions of U.S. dollars)
Brazil	4,844.80	reais	2.157	2,246.00
Canada	1,881.20	dollars	1.030	1,826.80
China	58,667.30	yuan	6.196	9,469.10

Country	GDP in Billions of Domestic Currency		Domestic Currency/U.S. Dollars (PPP Equivalent)	GDP (in billions of U.S. dollars)
Egypt	1,753.30	pounds	6.460	271.40
Germany	2,737.60	euros	0.753	3,636.00
India	113,550.70	rupees	60.502	1,876.80
Japan	478,075.30	yen	97.596	4,898.50
Mexico	16,104.40	pesos	12.772	1,260.90
South Korea	1,428,294.70	won	1,094.925	1,304.467
United Kingdom	1,612.80	pounds	0.639	2,523.20
United States	16,768.10	dollars	1.000	16,768.10
Comparing GDPs Across Countries, 2013(Source: http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx)				

GDP Per Capita

The U.S. economy has the largest GDP in the world, by a considerable amount. The United States is also a populous country; in fact, it is the third largest country by population in the world, although well behind China and India. So is the U.S. economy larger than other countries just because the United States has more people than most other countries, or because the U.S. economy is actually larger on a per-person basis? This question can be answered by calculating a country's **GDP per capita**; that is, the GDP divided by the population.

Equation:

$$\text{GDP per capita} = \text{GDP} / \text{population}$$

The second column of [\[link\]](#) lists the GDP of the same selection of countries that appeared in the previous [Tracking Real GDP over Time](#) and [\[link\]](#), showing their GDP as converted into U.S. dollars (which is the same as the last column of the previous table). The third column gives the population for each country. The fourth column lists the GDP per capita. GDP per capita is obtained in two steps: First, by dividing column two (GDP, in billions of dollars) by 1000 so it has the same units as column three (Population, in millions). Then dividing the result (GDP in millions of dollars) by column three (Population, in millions).

Country	GDP (in billions of U.S. dollars)	Population (in millions)	Per Capita GDP (in U.S. dollars)
Brazil	2,246.00	199.20	11,172.50
Canada	1,826.80	35.10	52,037.10
China	9,469.10	1,360.80	6,958.70
Egypt	271.40	83.70	3,242.90
Germany	3,636.00	80.80	44,999.50
India	1,876.80	1,243.30	1,509.50
Japan	4,898.50	127.3	38,467.80
Mexico	1,260.90	118.40	10,649.90
South Korea	1,304.47	50.20	25,975.10

Country	GDP (in billions of U.S. dollars)	Population (in millions)	Per Capita GDP (in U.S. dollars)
United Kingdom	2,523.20	64.10	39,371.70
United States	16,768.10	316.30	53,001.00

GDP Per Capita, 2013(Source: <http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx>)

Notice that the ranking by GDP is different from the ranking by GDP per capita. India has a somewhat larger GDP than Germany, but on a per capita basis, Germany has more than 10 times India's standard of living. Will China soon have a better standard of living than the U.S.? Read the following Clear It Up feature to find out.

Note:

Is China going to surpass the United States in terms of standard of living? As shown in [\[link\]](#), China has the second largest GDP of the countries: \$9.5 trillion compared to the United States' \$16.8 trillion. Perhaps it will surpass the United States, but probably not any time soon. China has a much larger population so that in per capita terms, its GDP is less than one fifth that of the United States (\$6,958.70 compared to \$53,001). The Chinese people are still quite poor relative to the United States and other developed countries. One caveat: For reasons to be discussed shortly, GDP per capita can give us only a rough idea of the differences in living standards across countries.

The high-income nations of the world—including the United States, Canada, the Western European countries, and Japan—typically have GDP per capita in the range of \$20,000 to \$50,000. Middle-income countries, which include much of Latin America, Eastern Europe, and some countries in East Asia, have GDP per capita in the range of \$6,000 to \$12,000. The low-income countries in the world, many of them located in Africa and Asia, often have GDP per capita of less than \$2,000 per year.

Key Concepts and Summary

Since GDP is measured in a country's currency, in order to compare different countries' GDPs, we need to convert them to a common currency. One way to do that is with the exchange rate, which is the price of one country's currency in terms of another. Once GDPs are expressed in a common currency, we can compare each country's GDP per capita by dividing GDP by population. Countries with large populations often have large GDPs, but GDP alone can be a misleading indicator of the wealth of a nation. A better measure is GDP per capita.

Self-Check Question

Exercise:

Problem:

Is it possible for GDP to rise while at the same time per capita GDP is falling? Is it possible for GDP to fall while per capita GDP is rising?

Solution:

Yes. The answer to both questions depends on whether GDP is growing faster or slower than population. If population grows faster than GDP, GDP increases, while GDP per capita decreases. If GDP falls, but population falls faster, then GDP decreases, while GDP per capita increases.

Exercise:

Problem:

The Central African Republic has a GDP of 1,107,689 million CFA francs and a population of 4.862 million. The exchange rate is 284.681 CFA francs per dollar. Calculate the GDP per capita of Central African Republic.

Solution:

Start with Central African Republic's GDP measured in francs. Divide it by the exchange rate to convert to U.S. dollars, and then divide by population to obtain the per capita figure. That is, $1,107,689 \text{ million francs} / 284.681 \text{ francs per dollar} / 4.862 \text{ million people} = \$800.28 \text{ GDP per capita}$.

Review Question

Exercise:

Problem:

What are the two main difficulties that arise in comparing the GDP of different countries?

Critical Thinking Question

Exercise:

Problem:

Cross country comparisons of GDP per capita typically use purchasing power parity equivalent exchange rates, which are a measure of the long run equilibrium value of an exchange rate. In fact, we used PPP equivalent exchange rates in this module. Why could using market exchange rates, which sometimes change dramatically in a short period of time, be misleading?

Exercise:

Problem:

Why might per capita GDP be only an imperfect measure of a country's standard of living?

Problems

Exercise:

Problem:

Ethiopia has a GDP of \$8 billion (measured in U.S. dollars) and a population of 55 million. Costa Rica has a GDP of \$9 billion (measured in U.S. dollars) and a population of 4 million. Calculate the per capita GDP for each country and identify which one is higher.

Exercise:

Problem:

In 1980, Denmark had a GDP of \$70 billion (measured in U.S. dollars) and a population of 5.1 million. In 2000, Denmark had a GDP of \$160 billion (measured in U.S. dollars) and a population of 5.3 million. By what percentage did Denmark's GDP per capita rise between 1980 and 2000?

Exercise:**Problem:**

The Czech Republic has a GDP of 1,800 billion koruny. The exchange rate is 20 koruny/U.S. dollar. The Czech population is 20 million. What is the GDP per capita of the Czech Republic expressed in U.S. dollars?

Glossary

exchange rate

the price of one currency in terms of another currency

GDP per capita

GDP divided by the population

How Well GDP Measures the Well-Being of Society

By the end of this section, you will be able to:

- Discuss how productivity influences the standard of living
- Explain the limitations of GDP as a measure of the standard of living
- Analyze the relationship between GDP data and fluctuations in the standard of living

The level of GDP per capita clearly captures some of what we mean by the phrase “standard of living.” Most of the migration in the world, for example, involves people who are moving from countries with relatively low GDP per capita to countries with relatively high GDP per capita.

“Standard of living” is a broader term than GDP. While GDP focuses on production that is bought and sold in markets, **standard of living** includes all elements that affect people’s well-being, whether they are bought and sold in the market or not. To illuminate the gap between GDP and standard of living, it is useful to spell out some things that GDP does not cover that are clearly relevant to standard of living.

Limitations of GDP as a Measure of the Standard of Living

While GDP includes spending on recreation and travel, it does not cover leisure time. Clearly, however, there is a substantial difference between an economy that is large because people work long hours, and an economy that is just as large because people are more productive with their time so they do not have to work as many hours. The GDP per capita of the U.S. economy is larger than the GDP per capita of Germany, as was shown in [\[link\]](#), but does that prove that the standard of living in the United States is higher? Not necessarily, since it is also true that the average U.S. worker works several hundred hours more per year more than the average German worker. The calculation of GDP does not take the German worker’s extra weeks of vacation into account.

While GDP includes what is spent on environmental protection, healthcare, and education, it does not include actual levels of environmental

cleanliness, health, and learning. GDP includes the cost of buying pollution-control equipment, but it does not address whether the air and water are actually cleaner or dirtier. GDP includes spending on medical care, but does not address whether life expectancy or infant mortality have risen or fallen. Similarly, it counts spending on education, but does not address directly how much of the population can read, write, or do basic mathematics.

GDP includes production that is exchanged in the market, but it does not cover production that is not exchanged in the market. For example, hiring someone to mow your lawn or clean your house is part of GDP, but doing these tasks yourself is not part of GDP. One remarkable change in the U.S. economy in recent decades is that, as of 1970, only about 42% of women participated in the paid labor force. By the second decade of the 2000s, nearly 60% of women participated in the paid labor force according to the Bureau of Labor Statistics. As women are now in the labor force, many of the services they used to produce in the non-market economy like food preparation and child care have shifted to some extent into the market economy, which makes the GDP appear larger even if more services are not actually being consumed.

GDP has nothing to say about the level of inequality in society. GDP per capita is only an average. When GDP per capita rises by 5%, it could mean that GDP for everyone in the society has risen by 5%, or that of some groups has risen by more while that of others has risen by less—or even declined. GDP also has nothing in particular to say about the amount of variety available. If a family buys 100 loaves of bread in a year, GDP does not care whether they are all white bread, or whether the family can choose from wheat, rye, pumpernickel, and many others—it just looks at whether the total amount spent on bread is the same.

Likewise, GDP has nothing much to say about what technology and products are available. The standard of living in, for example, 1950 or 1900 was not affected only by how much money people had—it was also affected by what they could buy. No matter how much money you had in 1950, you could not buy an iPhone or a personal computer.

In certain cases, it is not clear that a rise in GDP is even a good thing. If a city is wrecked by a hurricane, and then experiences a surge of rebuilding

construction activity, it would be peculiar to claim that the hurricane was therefore economically beneficial. If people are led by a rising fear of crime, to pay for installation of bars and burglar alarms on all their windows, it is hard to believe that this increase in GDP has made them better off. In that same vein, some people would argue that sales of certain goods, like pornography or extremely violent movies, do not represent a gain to society's standard of living.

Does a Rise in GDP Overstate or Understate the Rise in the Standard of Living?

The fact that GDP per capita does not fully capture the broader idea of standard of living has led to a concern that the increases in GDP over time are illusory. It is theoretically possible that while GDP is rising, the standard of living could be falling if human health, environmental cleanliness, and other factors that are not included in GDP are worsening. Fortunately, this fear appears to be overstated.

In some ways, the rise in GDP understates the actual rise in the standard of living. For example, the typical workweek for a U.S. worker has fallen over the last century from about 60 hours per week to less than 40 hours per week. Life expectancy and health have risen dramatically, and so has the average level of education. Since 1970, the air and water in the United States have generally been getting cleaner. New technologies have been developed for entertainment, travel, information, and health. A much wider variety of basic products like food and clothing is available today than several decades ago. Because GDP does not capture leisure, health, a cleaner environment, the possibilities created by new technology, or an increase in variety, the actual rise in the standard of living for Americans in recent decades has exceeded the rise in GDP.

On the other side, rates of crime, levels of traffic congestion, and inequality of incomes are higher in the United States now than they were in the 1960s. Moreover, a substantial number of services that used to be provided, primarily by women, in the non-market economy are now part of the market economy that is counted by GDP. By ignoring these factors, GDP would tend to overstate the true rise in the standard of living.

Note:

Visit this [website](#) to read about the American Dream and standards of living.



GDP is Rough, but Useful

A high level of GDP should not be the only goal of macroeconomic policy, or government policy more broadly. Even though GDP does not measure the broader standard of living with any precision, it does measure production well and it does indicate when a country is materially better or worse off in terms of jobs and incomes. In most countries, a significantly higher GDP per capita occurs hand in hand with other improvements in everyday life along many dimensions, like education, health, and environmental protection.

No single number can capture all the elements of a term as broad as “standard of living.” Nonetheless, GDP per capita is a reasonable, rough-and-ready measure of the standard of living.

Note:

How is the Economy Doing? How Does One Tell?

To determine the state of the economy, one needs to examine economic indicators, such as GDP. To calculate GDP is quite an undertaking. It is the broadest measure of a nation’s economic activity and we owe a debt to Simon Kuznets, the creator of the measurement, for that.

The sheer size of the U.S. economy as measured by GDP is huge—as of the third quarter of 2013, \$16.6 trillion worth of goods and services were produced annually. Real GDP informed us that the recession of 2008–2009 was a severe one and that the recovery from that has been slow, but is improving. GDP per capita gives a rough estimate of a nation’s standard of living. This chapter is the building block for other chapters that explore more economic indicators such as unemployment, inflation, or interest rates, and perhaps more importantly, will explain how they are related and what causes them to rise or fall.

Key Concepts and Summary

GDP is an indicator of a society’s standard of living, but it is only a rough indicator. GDP does not directly take account of leisure, environmental quality, levels of health and education, activities conducted outside the market, changes in inequality of income, increases in variety, increases in technology, or the (positive or negative) value that society may place on certain types of output.

Self-Check Question

Exercise:

Problem:

Explain briefly whether each of the following would cause GDP to overstate or understate the degree of change in the broad standard of living.

- a. The environment becomes dirtier
- b. The crime rate declines
- c. A greater variety of goods become available to consumers
- d. Infant mortality declines

Solution:

- a. A dirtier environment would reduce the broad standard of living, but not be counted in GDP, so a rise in GDP would overstate the standard of living.
- b. A lower crime rate would raise the broad standard of living, but not be counted directly in GDP, and so a rise in GDP would understate the standard of living.
- c. A greater variety of goods would raise the broad standard of living, but not be counted directly in GDP, and so a rise in GDP would understate the rise in the standard of living.
- d. A decline in infant mortality would raise the broad standard of living, but not be counted directly in GDP, and so a rise in GDP would understate the rise in the standard of living.

Review Question

Exercise:

Problem:

List some of the reasons why GDP should not be considered an effective measure of the standard of living in a country.

Critical Thinking Questions

Exercise:

Problem: How might a “green” GDP be measured?

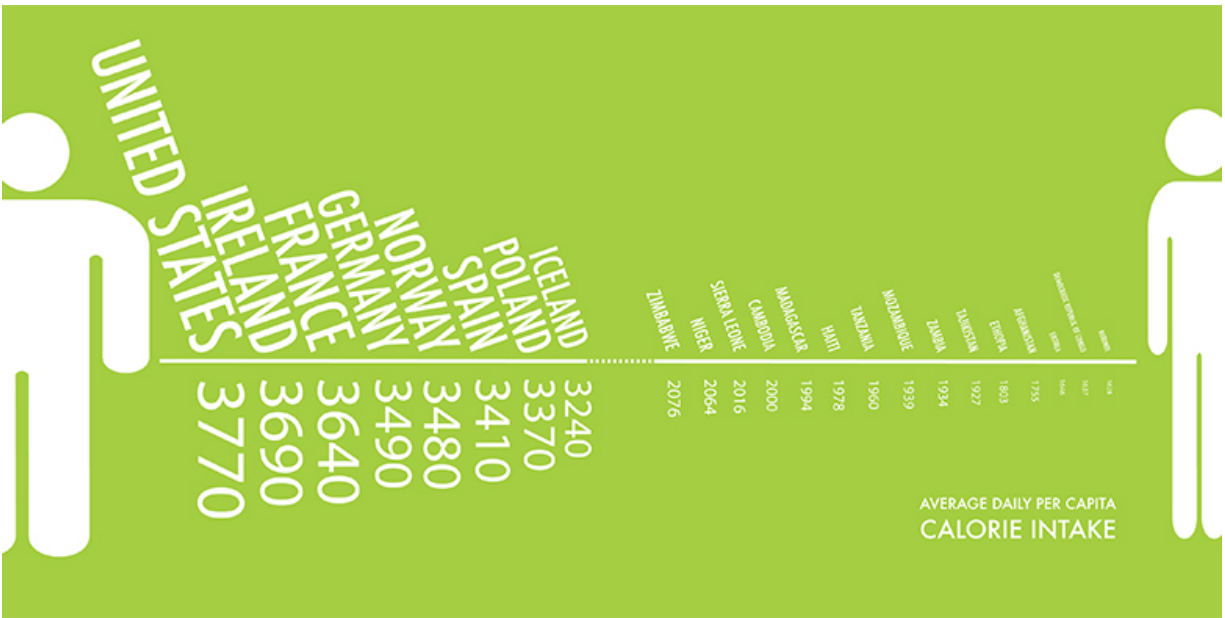
Glossary

standard of living

all elements that affect people’s happiness, whether these elements are bought and sold in the market or not

Introduction to Economic Growth
class="introduction"
Average Daily Calorie Consumption

Not only has
the number of
calories
consumed per
day increased,
so has the
amount of food
calories that
people are able
to afford based
on their
working
wages. (Credit:
modification of
work by
Lauren
Manning/Flick
r Creative
Commons)



Note:

Calories and Economic Growth

On average, humans need about 2,500 calories a day to survive, depending on height, weight, and gender. The economist Brad DeLong estimates that the average worker in the early 1600s earned wages that could afford him 2,500 food calories. This worker lived in Western Europe. Two hundred years later, that same worker could afford 3,000 food calories. However, between 1800 and 1875, just a time span of just 75 years, economic growth was so rapid that western European workers could purchase 5,000 food calories a day. By 2012, a low skilled worker in an affluent Western European/North American country could afford to purchase 2.4 million food calories per day.

What caused such a rapid rise in living standards between 1800 and 1875 and thereafter? Why is it that many countries, especially those in Western Europe, North America, and parts of East Asia, can feed their populations more than adequately, while others cannot? We will look at these and other questions as we examine long-run economic growth.

Note:

Introduction to Economic Growth

In this chapter, you will learn about:

- The Relatively Recent Arrival of Economic Growth
- Labor Productivity and Economic Growth
- Components of Economic Growth
- Economic Convergence

Every country worries about economic growth. In the United States and other high-income countries, the question is whether economic growth continues to provide the same remarkable gains in our standard of living as it did during the twentieth century. Meanwhile, can middle-income countries like South Korea, Brazil, Egypt, or Poland catch up to the higher-income countries? Or must they remain in the second tier of per capita income? Of the world's population of roughly 6.7 billion people, about 2.6 billion are scraping by on incomes that average less than \$2 per day, not that different from the standard of living 2,000 years ago. Can the world's poor be lifted from their fearful poverty? As the 1995 Nobel laureate in economics, Robert E. Lucas Jr., once noted: "The consequences for human welfare involved in questions like these are simply staggering: Once one starts to think about them, it is hard to think about anything else."

Dramatic improvements in a nation's standard of living are possible. After the Korean War in the late 1950s, the Republic of Korea, often called South Korea, was one of the poorest economies in the world. Most South Koreans worked in peasant agriculture. According to the British economist Angus Maddison, whose life's work was the measurement of GDP and population in the world economy, GDP per capita in 1990 international dollars was \$854 per year. From the 1960s to the early twenty-first century, a time period well within the lifetime and memory of many adults, the South Korean economy grew rapidly. Over these four decades, GDP per capita increased by more than 6% per year. According to the World Bank, GDP for South Korea now exceeds \$30,000 in nominal terms, placing it firmly among high-income countries like Italy, New Zealand, and Israel. Measured by total GDP in 2012, South Korea is the thirteenth-largest economy in the

world. For a nation of 49 million people, this transformation is extraordinary.

South Korea is a standout example, but it is not the only case of rapid and sustained economic growth. Other nations of East Asia, like Thailand and Indonesia, have seen very rapid growth as well. China has grown enormously since market-oriented economic reforms were enacted around 1980. GDP per capita in high-income economies like the United States also has grown dramatically albeit over a longer time frame. Since the Civil War, the U.S. economy has been transformed from a primarily rural and agricultural economy to an economy based on services, manufacturing, and technology.

The Relatively Recent Arrival of Economic Growth

By the end of this section, you will be able to:

- Explain the conditions that have allowed for modern economic growth in the last two centuries
- Analyze the influence of public policies on the long-run economic growth of an economy

Let's begin with a brief overview of the spectacular patterns of economic growth around the world in the last two centuries, commonly referred to as the period of **modern economic growth**. (Later in the chapter we will discuss lower rates of economic growth and some key ingredients for economic progress.) Rapid and sustained economic growth is a relatively recent experience for the human race. Before the last two centuries, although rulers, nobles, and conquerors could afford some extravagances and although economies rose above the subsistence level, the average person's standard of living had not changed much for centuries.

Progressive, powerful economic and institutional changes started to have a significant effect in the late eighteenth and early nineteenth centuries. According to the Dutch economic historian Jan Luiten van Zanden, slavery-based societies, favorable demographics, global trading routes, and standardized trading institutions that spread with different empires set the stage for the Industrial Revolution to succeed. The **Industrial Revolution** refers to the widespread use of power-driven machinery and the economic and social changes that resulted in the first half of the 1800s. Ingenious machines—the steam engine, the power loom, and the steam locomotive—performed tasks that otherwise would have taken vast numbers of workers to do. The Industrial Revolution began in Great Britain, and soon spread to the United States, Germany, and other countries.

The jobs for ordinary people working with these machines were often dirty and dangerous by modern standards, but the alternative jobs of that time in peasant agriculture and small-village industry were often dirty and dangerous, too. The new jobs of the Industrial Revolution typically offered higher pay and a chance for social mobility. A self-reinforcing cycle began: New inventions and investments generated profits, the profits provided funds for new investment and inventions, and the investments and inventions provided opportunities for further profits. Slowly, a group of national economies in Europe and North America emerged from centuries of sluggishness into a period of rapid modern growth. During the last two centuries, the average rate of growth of GDP per capita in the leading industrialized countries has averaged about 2% per year. What were times like before then? Read the following Clear It Up feature for the answer.

Note:

What were economic conditions like before 1870?

Angus Maddison, a quantitative economic historian, led the most systematic inquiry into national incomes before 1870. His methods recently have been refined and used to compile GDP per capita estimates from year 1 C.E. to 1348. [\[link\]](#) is an important counterpoint to most of the narrative in this chapter. It shows that nations can decline as well as rise. The declines in income are explained by a wide array of forces, such as epidemics, natural and weather-related disasters, the inability to govern large empires, and the remarkably slow pace of technological and institutional progress. Institutions are the traditions, laws, and so on by which people in a community agree to behave and govern themselves. Such institutions include marriage, religion, education, and laws of governance. Institutional progress is the development and codification of these institutions to reinforce social order, and thus, economic growth.

One example of such an institution is the Magna Carta (Great Charter), which the English nobles forced King John to sign in 1215. The Magna Carta codified the principles of due process, whereby a free man could not be penalized unless his peers had made a lawful judgment against him. This concept was later adopted by the United States in its own constitution. This social order may have contributed to England's GDP per capita in 1348, which was second to that of northern Italy.

In the study of economic growth, a country's institutional framework plays a critical role. [\[link\]](#) also shows relative global equality for almost 1,300 years. After this, we begin to see significant divergence in income (not shown in table).

Year	Northern Italy	Spain	England	Holland	Byzantium	Iraq	Egypt	Japan
1	\$800	\$600	\$600	\$600	\$700	\$700	\$700	-
730	-	-	-	-	-	\$920	\$730	\$402
1000	-	-	-	-	\$600	\$820	\$600	-
1150	-	-	-	-	\$580	\$680	\$660	\$520
1280	-	-	-	-	-	-	\$670	\$527
1300	\$1,588	\$864	\$892	-	-	-	\$610	-
1348	\$1,486	\$907	\$919	-	-	-	-	-

GDP Per Capita Estimates in Current International Dollars from AD 1 to 1348(Source: Bolt and van Zanden. “The First Update of the Maddison Project. Re-Estimating Growth Before 1820.” 2013)

Another fascinating and underreported fact is the high levels of income, compared to others at that time, attained by the Islamic Empire Abbasid Caliphate—which was founded in present-day Iraq in 730 C.E. At its height, the empire spanned large regions of the Middle East, North Africa, and Spain until its gradual decline over 200 years.

The Industrial Revolution led to increasing inequality among nations. Some economies took off, whereas others, like many of those in Africa or Asia, remained close to a subsistence standard of living. General calculations show that the 17 countries of the world with the most-developed economies had, on average, 2.4 times the GDP per capita of the world’s poorest economies in 1870. By 1960, the most developed economies had 4.2 times the GDP per capita of the poorest economies.

However, by the middle of the twentieth century, some countries had shown that catching up was possible. Japan’s economic growth took off in the 1960s and 1970s, with a growth rate of real GDP per capita averaging 11% per year during those decades. Certain countries in Latin America experienced a boom in economic growth in the 1960s as well. In Brazil, for example, GDP per capita expanded by an average annual rate of 11.1% from 1968 to 1973. In the 1970s, some East Asian economies, including South Korea, Thailand, and Taiwan, saw rapid growth. In these countries, growth rates of 11% to 12% per year in GDP per capita were not uncommon. More recently, China, with its population of 1.3 billion people, grew at a per capita rate 9% per year from 1984 into the 2000s. India, with a population of 1.1 billion, has shown promising signs of economic growth, with growth in GDP per capita of about 4% per year during the 1990s and climbing toward 7% to 8% per year in the 2000s.

Note:

Visit this [website](#) to read about the Asian Development Bank.



These waves of catch-up economic growth have not reached all shores. In certain African countries like Niger, Tanzania, and Sudan, for example, GDP per capita at the start of the 2000s was still less than \$300, not much higher than it was in the nineteenth century and for centuries before that. In the context of the overall situation of low-income people around the world, the good economic news from China (population: 1.3 billion) and India (population: 1.1 billion) is, nonetheless, astounding and heartening.

Economic growth in the last two centuries has made a striking change in the human condition. Richard Easterlin, an economist at the University of Southern California, wrote in 2000:

"By many measures, a revolution in the human condition is sweeping the world. Most people today are better fed, clothed, and housed than their predecessors two centuries ago. They are healthier, live longer, and are better educated. Women's lives are less centered on reproduction and political democracy has gained a foothold. Although Western Europe and its offshoots have been the leaders of this advance, most of the less developed nations have joined in during the 20th century, with the newly emerging nations of sub-Saharan Africa the latest to participate. Although the picture is not one of universal progress, it is the greatest advance in the human condition of the world's population ever achieved in such a brief span of time."

Rule of Law and Economic Growth

Economic growth depends on many factors. Key among those factors is adherence to the **rule of law** and protection of property rights and **contractual rights** by a country's government so that markets can work effectively and efficiently. Laws must be clear, public, fair, enforced, and equally applicable to all members of society. Property rights, as you might recall from [Environmental Protection and Negative Externalities](#) are the rights of individuals and firms to own property and use it as they see fit. If you have \$100, you have the right to use that money, whether you spend it, lend it, or keep it in a jar. It is your property. The definition of property includes physical property as well as the right to your training and experience, especially since your training is what determines your livelihood. The use of this property includes the right to enter into contracts with other parties with your property. Individuals or firms must own the property to enter into a contract.

Contractual rights, then, are based on property rights and they allow individuals to enter into agreements with others regarding the use of their property providing recourse through the legal system in the event of noncompliance. One example is the employment agreement: a skilled surgeon operates on an ill person and expects to get paid. Failure to pay would constitute a theft of property by the patient; that property being the services provided by the surgeon. In a society with strong property rights and contractual rights, the terms of the patient-surgeon contract will be fulfilled, because the surgeon would have recourse through the court system to extract payment from that individual. Without a legal system that enforces contracts, people would not be likely to enter into contracts for current or future services because of the risk of non-payment. This would make it difficult to transact business and would slow economic growth.

The World Bank considers a country's legal system effective if it upholds property rights and contractual rights. The World Bank has developed a ranking system for countries' legal systems based on effective protection of property rights and rule-based governance using a scale from 1 to 6, with 1 being the lowest and 6 the highest rating. In 2013, the world average ranking was 2.9. The three countries with the lowest ranking of 1.5 were Afghanistan, the Central African Republic, and Zimbabwe; their GDP per capita was \$679, \$333, and \$1,007 respectively. Afghanistan is cited by the World Bank as having a low standard of living, weak government structure, and lack of adherence to the rule of law, which has stymied its economic growth. The landlocked Central African Republic has poor economic resources as well as political instability and is a source of children used in human trafficking. Zimbabwe has had declining growth since 1998. Land redistribution and price controls have disrupted the economy, and corruption and violence have dominated the political process. Although global economic growth has increased, those countries lacking a clear system of property rights and an independent court system free from corruption have lagged far behind.

Key Concepts and Summary

Since the early nineteenth century, there has been a spectacular process of long-run economic growth during which the world's leading economies—mostly those in Western Europe and North America—expanded GDP per capita at

an average rate of about 2% per year. In the last half-century, countries like Japan, South Korea, and China have shown the potential to catch up. The extensive process of economic growth, often referred to as modern economic growth, was facilitated by the Industrial Revolution, which increased worker productivity and trade, as well as the development of governance and market institutions.

Self-Check Questions

Exercise:

Problem: Explain what the Industrial Revolution was and where it began.

Solution:

The Industrial Revolution refers to the widespread use of power-driven machinery and the economic and social changes that resulted in the first half of the 1800s. Ingenious machines—the steam engine, the power loom, and the steam locomotive—performed tasks that would have taken vast numbers of workers to do. The Industrial Revolution began in Great Britain, and soon spread to the United States, Germany, and other countries.

Exercise:

Problem:

Explain the difference between property rights and contractual rights. Why do they matter to economic growth?

Solution:

Property rights are the rights of individuals and firms to own property and use it as they see fit. Contractual rights are based on property rights and they allow individuals to enter into agreements with others regarding the use of their property providing recourse through the legal system in the event of noncompliance. Economic growth occurs when the standard of living increases in an economy, which occurs when output is increasing and incomes are rising. For this to happen, societies must create a legal environment that gives individuals the ability to use their property to their fullest and highest use, including the right to trade or sell that property. Without a legal system that enforces contracts, people would not be likely to enter into contracts for current or future services because of the risk of non-payment. This would make it difficult to transact business and would slow economic growth.

Review Questions

Exercise:

Problem:

How did the Industrial Revolution increase the rate of economic growth and income levels in the United States?

Exercise:

Problem:

How much should a nation be concerned if its rate of economic growth is just 2% slower than other nations?

Critical Thinking Question

Exercise:

Problem:

Over the past 50 years, many countries have experienced an annual growth rate in real GDP per capita greater than that of the United States. Some examples are China, Japan, South Korea, and Taiwan. Does that mean the United States is regressing relative to other countries? Does that mean these countries will eventually overtake the United States in terms of rate of growth of real GDP per capita? Explain.

References

Bolt, Jutta, and Jan Luiten van Zanden. "The Maddison Project: The First Update of the Maddison Project Re-Estimating Growth Before 1820 (Maddison-Project Working Paper WP-4)." *University of Groningen: Groningen Growth and Development Centre*. Last modified January 2013. <http://www.ggdc.net/maddison/publications/pdf/wp4.pdf>.

Central Intelligence Agency. "The World Factbook: Country Comparison GDP (Purchasing Power Parity)." <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html>.

DeLong, Brad. "Lighting the Rocket of Growth and Lightening the Toil of Work: Another Outtake from My 'Slouching Towards Utopia' MS...." *This is Brad DeLong's Grasping Reality* (blog). September 3, 2013. <http://delong.typepad.com/sdj/2013/09/lighting-the-rocket-of-growth-and-lightening-the-toil-of-work-another-outtake-from-my-slouching-towards-utopia-ms.html>.

Easterlin, Richard A. "The Worldwide Standard of Living since 1800." *The Journal of Economic Perspectives*. no. 1 (2000): 7–26. <http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.14.1.7>.

Maddison, Angus. *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History*. Oxford: Oxford University Press, 2007.

British Library. "Treasures in Full: Magna Carta." <http://www.bl.uk/treasures/magnacarta/>.

Rothbard, Murray N. Ludwig von Mises Institute. "Property Rights and the Theory of Contracts." *The Ethics of Liberty*. Last modified June 22, 2007. <http://mises.org/daily/2580>.

Salois, Matthew J., J. Richard Tiffin, and Kelvin George Balcombe. IDEAS: Research Division of the Federal Reserve Bank of St. Louis. "Impact of Income on Calorie and Nutrient Intakes: A Cross-Country Analysis." Presentation at the annual meeting of the Agricultural and Applied Economics Association, Pittsburgh, PA, July 24–26, 2011. <http://ideas.repec.org/p/ags/aaea11/103647.html>.

van Zanden, Jan Luiten. *The Long Road to the Industrial Revolution: The European Economy in a Global Perspective, 1000–1800 (Global Economic History Series)*. Boston: Brill, 2009.

The World Bank. "CPIA Property Rights and Rule-based Governance Rating (1=low to 6=high)." <http://data.worldbank.org/indicator/IQ.CPA.PROP.XQ>.

Rex A. Hudson, ed. *Brazil: A Country Study*. "Spectacular Growth, 1968–73." Washington: GPO for the Library of Congress, 1997. <http://countrystudies.us/brazil/64.htm>.

Glossary

contractual rights

the rights of individuals to enter into agreements with others regarding the use of their property providing recourse through the legal system in the event of noncompliance

Industrial Revolution

the widespread use of power-driven machinery and the economic and social changes that occurred in the first half of the 1800s

modern economic growth

the period of rapid economic growth from 1870 onward

rule of law

the process of enacting laws that protect individual and entity rights to use their property as they see fit. Laws must be clear, public, fair, and enforced, and applicable to all members of society

Components of Economic Growth

By the end of this section, you will be able to:

- Discuss the components of economic growth, including physical capital, human capital, and technology
- Explain capital deepening and its significance
- Analyze the methods employed in economic growth accounting studies
- Identify factors that contribute to a healthy climate for economic growth

Over decades and generations, seemingly small differences of a few percentage points in the annual rate of economic growth make an enormous difference in GDP per capita. In this module, we discuss some of the components of economic growth, including physical capital, human capital, and technology.

The category of **physical capital** includes the plant and equipment used by firms and also things like roads (also called **infrastructure**). Again, greater physical capital implies more output. Physical capital can affect productivity in two ways: (1) an increase in the *quantity* of physical capital (for example, more computers of the same quality); and (2) an increase in the *quality* of physical capital (same number of computers but the computers are faster, and so on). Human capital and physical capital accumulation are similar: In both cases, investment now pays off in longer-term productivity in the future.

The category of **technology** is the “joker in the deck.” Earlier we described it as the combination of invention and innovation. When most people think of new technology, the invention of new products like the laser, the smartphone, or some new wonder drug come to mind. In food production, the development of more drought-resistant seeds is another example of technology. Technology, as economists use the term, however, includes still more. It includes new ways of organizing work, like the invention of the assembly line, new methods for ensuring better quality of output in factories, and innovative institutions that facilitate the process of converting inputs into output. In short, technology comprises all the advances that

make the existing machines and other inputs produce more, and at higher quality, as well as altogether new products.

It may not make sense to compare the GDPs of China and say, Benin, simply because of the great difference in population size. To understand economic growth, which is really concerned with the growth in living standards of an average person, it is often useful to focus on GDP per capita. Using GDP per capita also makes it easier to compare countries with smaller numbers of people, like Belgium, Uruguay, or Zimbabwe, with countries that have larger populations, like the United States, the Russian Federation, or Nigeria.

To obtain a per capita production function, divide each input in [\[link\]](#)(a) by the population. This creates a second aggregate production function where the output is GDP per capita (that is, GDP divided by population). The inputs are the average level of human capital per person, the average level of physical capital per person, and the level of technology per person—see [\[link\]](#)(b). The result of having population in the denominator is mathematically appealing. Increases in population lower per capita income. However, increasing population is important for the average person only if the rate of income growth exceeds population growth. A more important reason for constructing a per capita production function is to understand the contribution of human and physical capital.

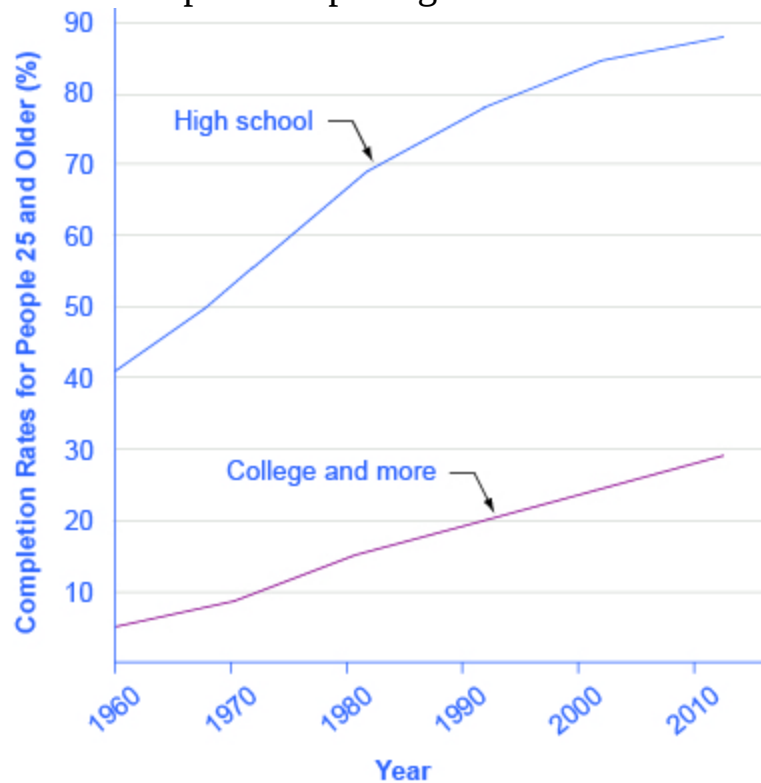
Capital Deepening

When society increases the level of capital per person, the result is called **capital deepening**. The idea of capital deepening can apply both to additional human capital per worker and to additional physical capital per worker.

Recall that one way to measure human capital is to look at the average levels of education in an economy. [\[link\]](#) illustrates the human capital deepening for U.S. workers by showing that the proportion of the U.S. population with a high school and a college degree is rising. As recently as 1970, for example, only about half of U.S. adults had at least a high school diploma; by the start of the twenty-first century, more than 80% of adults

had graduated from high school. The idea of human capital deepening also applies to the years of experience that workers have, but the average experience level of U.S. workers has not changed much in recent decades. Thus, the key dimension for deepening human capital in the U.S. economy focuses more on additional education and training than on a higher average level of work experience.

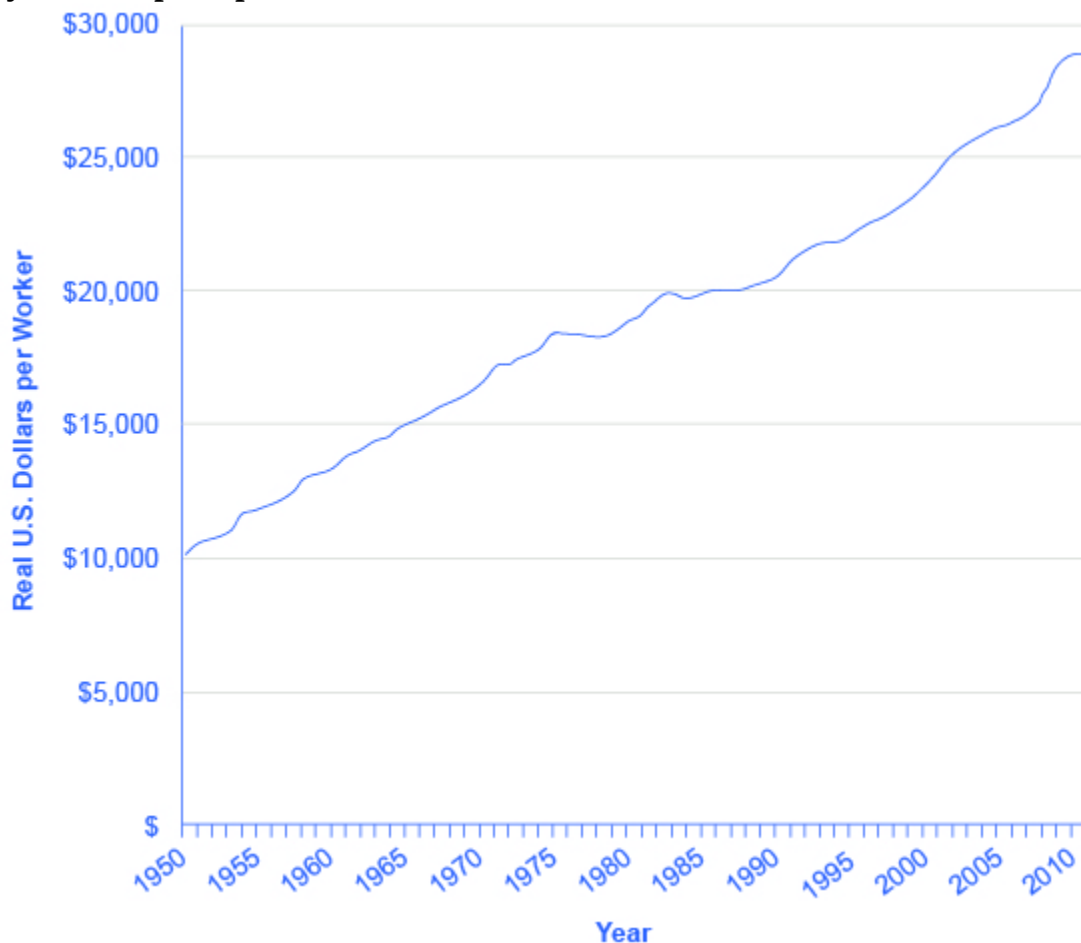
Human Capital Deepening in the U.S.



Rising levels of education for persons 25 and older show the deepening of human capital in the U.S. economy. Even today, relatively few U.S. adults have completed a four-year college degree. There is clearly room for additional deepening of human capital to occur. (Source: US Department of Education, National Center for Education Statistics)

Physical capital deepening in the U.S. economy is shown in [\[link\]](#). The average U.S. worker in the late 2000s was working with physical capital worth almost three times as much as that of the average worker of the early 1950s.

Physical Capital per Worker in the United States



The value of the physical capital, measured by plant and equipment, used by the average worker in the U.S. economy has risen over the decades. The increase may have leveled off a bit in the 1970s and 1980s, which were not, coincidentally, times of slower-than-usual growth in worker productivity. We see a renewed increase in physical capital per worker in the late 1990s, followed by a flattening in the early 2000s. (Source: Center for International Comparisons of Production, Income and Prices, University of Pennsylvania)

Not only does the current U.S. economy have better-educated workers with more and improved physical capital than it did several decades ago, but these workers have access to more advanced technologies. Growth in technology is impossible to measure with a simple line on a graph, but evidence that we live in an age of technological marvels is all around us—discoveries in genetics and in the structure of particles, the wireless Internet, and other inventions almost too numerous to count. The U.S. Patent and Trademark Office typically has issued more than 150,000 patents annually in recent years.

This recipe for economic growth—investing in labor productivity, with investments in human capital and technology, as well as increasing physical capital—also applies to other economies. In South Korea, for example, universal enrollment in primary school (the equivalent of kindergarten through sixth grade in the United States) had already been achieved by 1965, when Korea's GDP per capita was still near its rock bottom low. By the late 1980s, Korea had achieved almost universal secondary school education (the equivalent of a high school education in the United States). With regard to physical capital, Korea's rates of investment had been about 15% of GDP at the start of the 1960s, but doubled to 30–35% of GDP by the late 1960s and early 1970s. With regard to technology, South Korean students went to universities and colleges around the world to get scientific and technical training, and South Korean firms reached out to study and form partnerships with firms that could offer them technological insights. These factors combined to foster South Korea's high rate of economic growth.

Growth Accounting Studies

Since the late 1950s, economists have conducted growth accounting studies to determine the extent to which physical and human capital deepening and technology have contributed to growth. The usual approach uses an aggregate production function to estimate how much of per capita economic growth can be attributed to growth in physical capital and human capital. These two inputs can be measured, at least roughly. The part of growth that is unexplained by measured inputs, called the residual, is then attributed to

growth in technology. The exact numerical estimates differ from study to study and from country to country, depending on how researchers measured these three main factors over what time horizons. For studies of the U.S. economy, three lessons commonly emerge from growth accounting studies.

First, technology is typically the most important contributor to U.S. economic growth. Growth in human capital and physical capital often explains only half or less than half of the economic growth that occurs. New ways of doing things are tremendously important.

Second, while investment in physical capital is essential to growth in labor productivity and GDP per capita, building human capital is at least as important. Economic growth is not just a matter of more machines and buildings. One vivid example of the power of human capital and technological knowledge occurred in Europe in the years after World War II (1939–1945). During the war, a large share of Europe’s physical capital, such as factories, roads, and vehicles, was destroyed. Europe also lost an overwhelming amount of human capital in the form of millions of men, women, and children who died during the war. However, the powerful combination of skilled workers and technological knowledge, working within a market-oriented economic framework, rebuilt Europe’s productive capacity to an even higher level within less than two decades.

A third lesson is that these three factors of human capital, physical capital, and technology work together. Workers with a higher level of education and skills are often better at coming up with new technological innovations. These technological innovations are often ideas that cannot increase production until they become a part of new investment in physical capital. New machines that embody technological innovations often require additional training, which builds worker skills further. If the recipe for economic growth is to succeed, an economy needs all the ingredients of the aggregate production function. See the following Clear It Up feature for an example of how human capital, physical capital, and technology can combine to significantly impact lives.

Note:

How do girls' education and economic growth relate in low-income countries?

In the early 2000s, according to the World Bank, about 110 million children between the ages of 6 and 11 were not in school—and about two-thirds of them were girls. In Bangladesh, for example, the illiteracy rate for those aged 15 to 24 was 78% for females, compared to 75% for males. In Egypt, for this age group, illiteracy was 84% for females and 91% for males. Cambodia had 86% illiteracy for females and 88% for males. Nigeria had 66% illiteracy for females in the 15 to 24 age bracket and 78% for males.

Whenever any child does not receive a basic education, it is both a human and an economic loss. In low-income countries, wages typically increase by an average of 10 to 20% with each additional year of education. There is, however, some intriguing evidence that helping girls in low-income countries to close the education gap with boys may be especially important, because of the social role that many of the girls will play as mothers and homemakers.

Girls in low-income countries who receive more education tend to grow up to have fewer, healthier, better-educated children. Their children are more likely to be better nourished and to receive basic health care like immunizations. Economic research on women in low-income economies backs up these findings. When 20 women get one additional year of schooling, as a group they will, on average, have one less child. When 1,000 women get one additional year of schooling, on average one to two fewer women from that group will die in childbirth. When a woman stays in school an additional year, that factor alone means that, on average, each of her children will spend an additional half-year in school. Education for girls is a good investment because it is an investment in economic growth with benefits beyond the current generation.

A Healthy Climate for Economic Growth

While physical and human capital deepening and better technology are important, equally important to a nation's well-being is the climate or system within which these inputs are cultivated. Both the type of market

economy and a legal system that governs and sustains property rights and contractual rights are important contributors to a healthy economic climate.

A healthy economic climate usually involves some sort of market orientation at the microeconomic, individual, or firm decision-making level. Markets that allow personal and business rewards and incentives for increasing human and physical capital encourage overall macroeconomic growth. For example, when workers participate in a competitive and well-functioning labor market, they have an incentive to acquire additional human capital, because additional education and skills will pay off in higher wages. Firms have an incentive to invest in physical capital and in training workers, because they expect to earn higher profits for their shareholders. Both individuals and firms look for new technologies, because even small inventions can make work easier or lead to product improvement. Collectively, such individual and business decisions made within a market structure add up to macroeconomic growth. Much of the rapid growth since the late nineteenth century has come from harnessing the power of competitive markets to allocate resources. This market orientation typically reaches beyond national borders and includes openness to international trade.

A general orientation toward markets does not rule out important roles for government. There are times when markets fail to allocate capital or technology in a manner that provides the greatest benefit for society as a whole. The role of the government is to correct these failures. In addition, government can guide or influence markets toward certain outcomes. The following examples highlight some important areas that governments around the world have chosen to invest in to facilitate capital deepening and technology:

- Education. The Danish government requires all children under 16 to attend school. They can choose to attend a public school (*Folkeskole*) or a private school. Students do not pay tuition to attend *Folkeskole*. Thirteen percent of primary/secondary (elementary/high) school is private, and the government supplies vouchers to citizens who choose private school.

- **Savings and Investment.** In the United States, as in other countries, private investment is taxed. Low capital gains taxes encourage investment and so also economic growth.
- **Infrastructure.** The Japanese government in the mid-1990s undertook significant infrastructure projects to improve roads and public works. This in turn increased the stock of physical capital and ultimately economic growth.
- **Special Economic Zones.** The island of Mauritius is one of the few African nations to encourage international trade in government-supported **special economic zones (SEZ)**. These are areas of the country, usually with access to a port where, among other benefits, the government does not tax trade. As a result of its SEZ, Mauritius has enjoyed above-average economic growth since the 1980s. Free trade does not have to occur in an SEZ however. Governments can encourage international trade across the board, or surrender to protectionism.
- **Scientific Research.** The European Union has strong programs to invest in scientific research. The researchers Abraham García and Pierre Mohnen demonstrate that firms which received support from the Austrian government actually increased their research intensity and had more sales. Governments can support scientific research and technical training that helps to create and spread new technologies. Governments can also provide a legal environment that protects the ability of inventors to profit from their inventions.

There are many more ways in which the government can play an active role in promoting economic growth; we explore them in other chapters and in particular in [Macroeconomic Policy Around the World](#). A healthy climate for growth in GDP per capita and labor productivity includes human capital deepening, physical capital deepening, and technological gains, operating in a market-oriented economy with supportive government policies.

Key Concepts and Summary

Over decades and generations, seemingly small differences of a few percentage points in the annual rate of economic growth make an enormous difference in GDP per capita. Capital deepening refers to an increase in the

amount of capital per worker, either human capital per worker, in the form of higher education or skills, or physical capital per worker. Technology, in its economic meaning, refers broadly to all new methods of production, which includes major scientific inventions but also small inventions and even better forms of management or other types of institutions. A healthy climate for growth in GDP per capita consists of improvements in human capital, physical capital, and technology, in a market-oriented environment with supportive public policies and institutions.

Self-Check Questions

Exercise:

Problem:

What do the growth accounting studies conclude are the determinants of growth? Which is more important, the determinants or how they are combined?

Solution:

Capital deepening and technology are important. What seems to be more important is how they are combined.

Exercise:

Problem:

What policies can the government of a free-market economy implement to stimulate economic growth?

Solution:

Government can contribute to economic growth by investing in human capital through the education system, building a strong physical infrastructure for transportation and commerce, increasing investment by lowering capital gains taxes, creating special economic zones that allow for reduced tariffs, and investing in research and development.

Exercise:

Problem:

List the areas where government policy can help economic growth.

Solution:

Public education, low investment taxes, funding for infrastructure projects, special economic zones

Review Questions**Exercise:**

Problem: What is an aggregate production function?

Exercise:

Problem: What is capital deepening?

Exercise:**Problem:**

What do economists mean when they refer to improvements in technology?

Critical Thinking Questions**Exercise:****Problem:**

Education seems to be important for human capital deepening. As people become better educated and more knowledgeable, are there limits to how much additional benefit more education can provide? Why or why not?

Exercise:

Problem:

Describe some of the political and social tradeoffs that might occur when a less developed country adopts a strategy to promote labor force participation and economic growth via investment in girls' education.

Exercise:

Problem: Why is investing in girls' education beneficial for growth?

Exercise:**Problem:**

How is the concept of technology, as defined with the aggregate production function, different from our everyday use of the word?

References

"Women and the World Economy: A Guide to Womenomics." *The Economist*, April 12, 2006. <http://www.economist.com/node/6802551>.

Farole, Thomas, and Gokhan Akinci, eds. *Special Economic Zones: Progress, Emerging Challenges, and Future Directions*. Washington: The World Bank, 2011. http://publications.worldbank.org/index.php?main_page=product_info&products_id=24138.

Garcia, Abraham, and Pierre Mohnen. United Nations University, Maastricht Economic and Social Research and Training Centre on Innovation and Technology: UNU-MERIT. "Impact of Government Support on R&D and Innovation (Working Paper Series #2010-034)." <http://www.merit.unu.edu/publications/wppdf/2010/wp2010-034.pdf>.

Heston, Alan, Robert Summers, and Bettina Aten. "Penn World Table Version 7.1." *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania*. Last modified July 2012. https://pwt.sas.upenn.edu/php_site/pwt71/pwt71_form.php.

United States Department of Labor: Bureau of Labor Statistics. "Women at Work: A Visual Essay." *Monthly Labor Review*, October 2003, 45–50.
<http://www.bls.gov/opub/mlr/2003/10/ressum3.pdf>.

Glossary

capital deepening

an increase by society in the average level of physical and/or human capital per person

infrastructure

a component of physical capital such as roads, rail systems, and so on

physical capital

the plant and equipment used by firms in production; this includes infrastructure

special economic zone (SEZ)

area of a country, usually with access to a port where, among other benefits, the government does not tax trade

technology

all the ways in which existing inputs produce more or higher quality, as well as different and altogether new products

Introduction to Unemployment
class="introduction"
Out of Business

Borders was
one of the
many
companies
unable to
recover from
the economic
recession of
2008-2009.

(Credit:
modification
of work by
Luis Villa del
Campo/Flick
r Creative
Commons)



Note:

The Mysterious Case of the Missing Candidates

Nearly eight million U.S. jobs were lost during the Great Recession of 2008-2009, with unemployment peaking at 10% in October 2009, according to the Bureau of Labor Statistics (BLS). That is a huge number of positions gone. During the tepid recovery, some positions were added, but as of summer 2013, unemployment had remained persistently higher than the pre-recession rate of less than 5%. Some economists and policymakers worried the recovery would be “jobless.” With the economy growing, albeit slowly, why wasn’t the unemployment number falling? Why were firms not hiring?

Peter Cappelli, noted Wharton management professor and Director of Wharton’s Center for Human Resources, does not believe the job search process is akin to what he terms the “Home Depot” view of hiring. According to him, this view “basically says that filling a job is like replacing a part in a washing machine. You simply find someone who does the exact same job as that broken part, plug him or her into the washing

machine and that is it.” The job search, for both the prospective employee and the employer, is more complex than that.

In a hiring situation, employers hold all the cards. They write the job descriptions, determine the salaries, decide when and how to advertise positions, and set the controls on employment application screening software. Advertising for positions has increased as the economic recovery progresses, yet here’s the kicker: Employers say there are no applicants out there who meet their needs. While the unemployment rate is now below 6% as of the beginning of 2015, many economists and policymakers (including the Chair of the Federal Reserve, Janet Yellen) are still concerned about "slack" in the labor market. So the question arises: where are the job candidates?

That question leads us to the topic of this chapter—unemployment. What constitutes it? How is it measured? And if the economy is growing, why isn’t the pool of job openings growing along with it? Sounds like the economy has a case of “missing” candidates.

Note:**Introduction to Unemployment**

In this chapter, you will learn about:

- How the Unemployment Rate is Defined and Computed
- Patterns of Unemployment
- What Causes Changes in Unemployment over the Short Run
- What Causes Changes in Unemployment over the Long Run

Unemployment can be a terrible and wrenching life experience—like a serious automobile accident or a messy divorce—whose consequences can be fully understood only by someone who has gone through it. For unemployed individuals and their families, there is the day-to-day financial stress of not knowing where the next paycheck is coming from. There are painful adjustments, like watching your savings account dwindle, selling a car and buying a cheaper one, or moving to a less expensive place to live.

Even when the unemployed person finds a new job, it may pay less than the previous one. For many people, their job is an important part of their self worth. When unemployment separates people from the workforce, it can affect family relationships as well as mental and physical health.

The human costs of unemployment alone would justify making a low level of unemployment an important public policy priority. But unemployment also includes economic costs to the broader society. When millions of unemployed but willing workers cannot find jobs, an economic resource is going unused. An economy with high unemployment is like a company operating with a functional but unused factory. The opportunity cost of unemployment is the output that could have been produced by the unemployed workers.

This chapter will discuss how the unemployment rate is defined and computed. It will examine the patterns of unemployment over time, for the U.S. economy as a whole, for different demographic groups in the U.S. economy, and for other countries. It will then consider an economic explanation for unemployment, and how it explains the patterns of unemployment and suggests public policies for reducing it.

How the Unemployment Rate is Defined and Computed

By the end of this section, you will be able to:

- Calculate the labor force percentage and the unemployment rate
- Explain hidden unemployment and what it means to be in or out of the labor force
- Evaluate the collection and interpretation of unemployment data

Unemployment is typically described in newspaper or television reports as a percentage or a rate. A recent report might have said, for example, *from August 2009 to November 2009, the U.S. unemployment rate rose from 9.7% to 10.0%, but by June 2010, it had fallen to 9.5%*. At a glance, the changes between the percentages may seem small. But remember that the U.S. economy has about 155 million adults who either have jobs or are looking for them. A rise or fall of just 0.1% in the unemployment rate of 155 million potential workers translates into 155,000 people, which is roughly the total population of a city like Syracuse, New York, Brownsville, Texas, or Pasadena, California. Large rises in the unemployment rate mean large numbers of job losses. In November 2009, at the peak of the recession, about 15 million people were out of work. Even with the unemployment rate now at 5.5% as of February 2015, about 8 million people total are out of work.

Note:

The [Bureau of Labor Statistics](#) tracks and reports all data related to unemployment.



Who's In or Out of the Labor Force?

Should everyone without a job be counted as unemployed? Of course not. Children, for example, should not be counted as unemployed. Surely, the retired should not be counted as unemployed. Many full-time college students have only a part-time job, or no job at all, but it seems inappropriate to count them as suffering the pains of unemployment. Some people are not working because they are rearing children, ill, on vacation, or on parental leave.

The point is that the adult population is not just divided into employed and unemployed. A third group exists: people who do not have a job, and for some reason—retirement, looking after children, taking a voluntary break before a new job—are not interested in having a job, either. It also includes those who do want a job but have quit looking, often due to being discouraged by their inability to find suitable employment. Economists refer to this third group of those who are not working and not looking for work as **out of the labor force** or not in the labor force.

The U.S. unemployment rate, which is based on a monthly survey carried out by the U.S. Bureau of the Census, asks a series of questions to divide up the adult population into employed, unemployed, or not in the labor force. To be classified as unemployed, a person must be without a job, currently available to work, and actively looking for work in the previous four weeks. Thus, a person who does not have a job but who is not currently available to work or has not actively looked for work in the last four weeks is counted as out of the labor force.

Employed: currently working for pay

Unemployed: Out of work and actively looking for a job

Out of the labor force: Out of paid work and not actively looking for a job

Labor force: the number of employed plus the unemployed

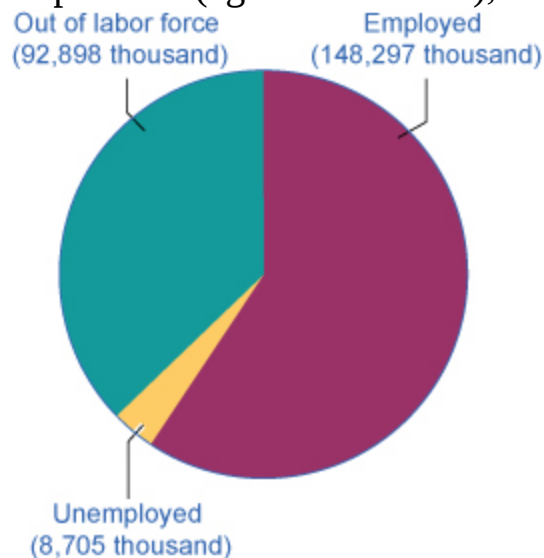
Calculating the Unemployment Rate

[\[link\]](#) shows the three-way division of the over-16 adult population. In February 2015, about 62.8% of the adult population was "in the labor force"; that is, people are either employed or without a job but looking for work. Those in the labor force can be divided into the employed and the unemployed. These values are also shown in [\[link\]](#). The **unemployment rate** is not the percentage of the total adult population without jobs, but rather the percentage of adults who are in the labor force but who do not have jobs:

Equation:

$$\text{Unemployment rate} = \frac{\text{Unemployed people}}{\text{Total labor force}} \times 100$$

Employed, Unemployed, and Out of the Labor Force Distribution of Adult Population (age 16 and older), February 2015



The total adult, working-age population in February 2015 was 249.9 million. Out of this total population, 148.3 were classified as employed, and 8.7 million were classified as unemployed. The remaining 92.9 were classified as out of the labor

force. As you will learn, however, this seemingly simple chart does not tell the whole story.

Total adult population over the age of 16	249.9 million
In the labor force	157 million (62.8%)
Employed	148.3 million
Unemployed	8.7 million
Out of the labor force	92.9 million (37.2%)

U.S. Employment and Unemployment, February 2015(Source: <http://www.bls.gov/news.release/empsit.t01.htm>)

In this example, the unemployment rate can be calculated as 8.7 million unemployed people divided by 157 million people in the labor force, which works out to a 5.5% rate of unemployment. The following Work It Out feature will walk you through the steps of this calculation.

Note:**Calculating Labor Force Percentages**

So how do economists arrive at the percentages in and out of the labor force and the unemployment rate? We will use the values in [\[link\]](#) to illustrate the steps.

To determine the percentage in the labor force:

Step 1. Divide the number of people in the labor force (157 million) by the total adult (working-age) population (249.9 million).

Step 2. Multiply by 100 to obtain the percentage.

Equation:

$$\begin{aligned}\text{Percentage in the labor force} &= \frac{157}{249.9} \\ &= 0.6282 \\ &= 62.8\%\end{aligned}$$

To determine the percentage out of the labor force:

Step 1. Divide the number of people out the labor force (92.9 million) by the total adult (working-age) population (249.9 million).

Step 2. Multiply by 100 to obtain the percentage.

Equation:

$$\begin{aligned}\text{Percentage in the labor force} &= \frac{92.9}{249.9} \\ &= 0.3717 \\ &= 37.2\%\end{aligned}$$

To determine the unemployment rate:

Step 1. Divide the number of unemployed people (8.7 million) by the total labor force (157 million).

Step 2. Multiply by 100 to obtain the rate.

Equation:

$$\begin{aligned}\text{Unemployment rate} &= \frac{8.7}{157} \\ &= 0.0554 \\ &= 5.5\%\end{aligned}$$

Hidden Unemployment

Even with the “out of the labor force” category, there are still some people that are mislabeled in the categorization of employed, unemployed, or out

of the labor force. There are some people who have only part time or temporary jobs and who are looking for full time and permanent employment that are counted as employed, though they are not employed in the way they would like or need to be. Additionally, there are individuals who are **underemployed**. This includes those that are trained or skilled for one type or level of work who are working in a lower paying job or one that does not utilize their skills. For example, an individual with a college degree in finance who is working as a sales clerk would be considered underemployed. They are, however, also counted in the employed group. All of these individuals fall under the umbrella of the term “hidden unemployment.” **Discouraged workers**, those who have stopped looking for employment and, hence, are no longer counted in the unemployed also fall into this group

Labor Force Participation Rate

Another important statistic is the **labor force participation rate**. This is the percentage of adults in an economy who are either employed or who are unemployed and looking for a job. So, using the data in [\[link\]](#) and [\[link\]](#), those included in this calculation would be the 157 million individuals in the labor force. The rate is calculated by taking the number of people in the labor force, that is, the number employed and the number unemployed, divided by the total adult population and multiplying by 100 to get the percentage. For the data from February 2015, the labor force participation rate is 62.8%. Historically, the civilian labor force participation rate in the United States climbed beginning in the 1960s as women increasingly entered the workforce, and it peaked at around 68% in late 1999 to early 2000. Since then, the labor force participation rate has steadily declined.

The Establishment Payroll Survey

When the unemployment report comes out each month, the Bureau of Labor Statistics (BLS) also reports on the number of jobs created—which comes from the establishment payroll survey. The payroll survey is based on a survey of about 140,000 businesses and government agencies throughout the United States. It generates payroll employment estimates by

the following criteria: all employees, average weekly hours worked, and average hourly, weekly, and overtime earnings. One of the criticisms of this survey is that it does not count the self-employed. It also does not make a distinction between new, minimum wage, part time or temporary jobs and full time jobs with “decent” pay.

How Is the U.S. Unemployment Data Collected?

The unemployment rate announced by the U.S. Bureau of Labor Statistics each month is based on the Current Population Survey (CPS), which has been carried out every month since 1940. Great care is taken to make this survey representative of the country as a whole. The country is first divided into 3,137 areas. The U.S. Bureau of the Census then selects 729 of these areas to survey. The 729 areas are then divided into districts of about 300 households each, and each district is divided into clusters of about four dwelling units. Every month, Census Bureau employees call about 15,000 of the four-household clusters, for a total of 60,000 households. Households are interviewed for four consecutive months, then rotated out of the survey for eight months, and then interviewed again for the same four months the following year, before leaving the sample permanently.

Based on this survey, unemployment rates are calculated by state, industry, urban and rural areas, gender, age, race or ethnicity, and level of education. A wide variety of other information is available, too. For example, how long have people been unemployed? Did they become unemployed because they quit, or were laid off, or their employer went out of business? Is the unemployed person the only wage earner in the family? The Current Population Survey is a treasure trove of information about employment and unemployment. If you are wondering what the difference is between the CPS and EPS, read the following Clear it Up feature.

Note:

What is the difference between CPS and EPS?

The Current Population Survey (CPS) conducted by the United States Census Bureau measures the percentage of the labor force that is

unemployed. The establishment payroll survey (EPS) by the Bureau of Labor Statistics is a payroll survey that measures the net change in jobs created for the month.

Criticisms of Measuring Unemployment

There are always complications in measuring the number of unemployed. For example, what about people who do not have jobs and would be available to work, but have gotten discouraged at the lack of available jobs in their area and stopped looking? Such people, and their families, may be suffering the pains of unemployment. But the survey counts them as out of the labor force because they are not actively looking for work. Other people may tell the Census Bureau that they are ready to work and looking for a job but, truly, they are not that eager to work and are not looking very hard at all. They are counted as unemployed, although they might more accurately be classified as out of the labor force. Still other people may have a job, perhaps doing something like yard work, child care, or cleaning houses, but are not reporting the income earned to the tax authorities. They may report being unemployed, when they actually are working.

Although the unemployment rate gets most of the public and media attention, economic researchers at the Bureau of Labor Statistics publish a wide array of surveys and reports that try to measure these kinds of issues and to develop a more nuanced and complete view of the labor market. It is not exactly a hot news flash that economic statistics are imperfect. Even imperfect measures like the unemployment rate, however, can still be quite informative, when interpreted knowledgeably and sensibly.

Note:

Click [here](#) to learn more about the CPS to read frequently asked questions about employment and labor.



Key Concepts and Summary

Unemployment imposes high costs. Unemployed individuals suffer from loss of income and from stress. An economy with high unemployment suffers an opportunity cost of unused resources. The adult population can be divided into those in the labor force and those out of the labor force. In turn, those in the labor force are divided into employed and unemployed. A person without a job must be willing and able to work and actively looking for work to be counted as unemployed; otherwise, a person without a job is counted as being out of the labor force. The unemployment rate is defined as the number of unemployed persons divided by the number of persons in the labor force (not the overall adult population). The Current Population Survey (CPS) conducted by the United States Census Bureau measures the percentage of the labor force that is unemployed. The establishment payroll survey by the Bureau of Labor Statistics measures the net change in jobs created for the month.

Self-Check Questions

Exercise:

Problem:

Suppose the adult population over the age of 16 is 237.8 million and the labor force is 153.9 million (of whom 139.1 million are employed). How many people are “not in the labor force?” What are the proportions of employed, unemployed and not in the labor force in the population? *Hint:* Proportions are percentages.

Solution:

The population is divided into those “in the labor force” and those “not in the labor force.” Thus, the number of adults not in the labor force is $237.8 - 153.9 = 83.9$ million. Since the labor force is divided into employed persons and unemployed persons, the number of unemployed persons is $153.9 - 139.1 = 14.8$ million. Thus, the adult population has the following proportions:

- $139.1/237.8 = 58.5\%$ employed persons
- $14.8/237.8 = 6.2\%$ unemployed persons
- $83.9/237.8 = 35.3\%$ persons out of the labor force

Exercise:**Problem:**

Using the above data, what is the unemployment rate? These data are U.S. statistics from 2010. How does it compare to the February 2015 unemployment rate computed earlier?

Solution:

The unemployment rate is defined as the number of unemployed persons as a percentage of the labor force or $14.8/153.9 = 9.6\%$. This is higher than the February 2015 unemployment rate, computed earlier, of 5.5%.

Review Questions**Exercise:****Problem:**

What is the difference between being unemployed and being out of the labor force?

Exercise:

Problem:

How is the unemployment rate calculated? How is the labor force participation rate calculated?

Exercise:

Problem: Are all adults who do not hold jobs counted as unemployed?

Exercise:

Problem:

If you are out of school but working part time, are you considered employed or unemployed in U.S. labor statistics? If you are a full time student and working 12 hours a week at the college cafeteria are you considered employed or not in the labor force? If you are a senior citizen who is collecting social security and a pension and working as a greeter at Wal-Mart are you considered employed or not in the labor force?

Exercise:

Problem:

What happens to the unemployment rate when unemployed workers are reclassified as discouraged workers?

Exercise:

Problem:

What happens to the labor force participation rate when employed individuals are reclassified as unemployed? What happens when they are reclassified as discouraged workers?

Exercise:

Problem:

What are some of the problems with using the unemployment rate as an accurate measure of overall joblessness?

Exercise:**Problem:**

What criteria are used by the BLS to count someone as employed? As unemployed?

Exercise:**Problem:**

Assess whether the following would be counted as “unemployed” in the Current Employment Statistics survey.

- A. A husband willingly stays home with children while his wife works.
- B. A manufacturing worker whose factory just closed down.
- C. A college student doing an unpaid summer internship.
- D. A retiree.
- E. Someone who has been out of work for two years but keeps looking for a job.
- F. Someone who has been out of work for two months but isn't looking for a job.
- G. Someone who hates her present job and is actively looking for another one.
- H. Someone who decides to take a part time job because she could not find a full time position.

Critical Thinking Questions**Exercise:****Problem:**

Using the definition of the unemployment rate, is an increase in the unemployment rate necessarily a bad thing for a nation?

Exercise:

Problem:

Is a decrease in the unemployment rate necessarily a good thing for a nation? Explain.

Exercise:**Problem:**

If many workers become discouraged from looking for jobs, explain how the number of jobs could decline but the unemployment rate could fall at the same time.

Exercise:**Problem:**

Would you expect hidden unemployment to be higher, lower, or about the same when the unemployment rate is high, say 10%, versus low, say 4%? Explain.

Problems**Exercise:****Problem:**

A country with a population of eight million adults has five million employed, 500,000 unemployed, and the rest of the adult population is out of the labor force. What's the unemployment rate? What share of population is in the labor force? Sketch a pie chart that divides the adult population into these three groups.

Glossary

discouraged workers

those who have stopped looking for employment due to the lack of suitable positions available

labor force participation rate

this is the percentage of adults in an economy who are either employed or who are unemployed and looking for a job

out of the labor force

those who are not working and not looking for work—whether they want employment or not; also termed “not in the labor force”

underemployed

individuals who are employed in a job that is below their skills

unemployment rate

the percentage of adults who are in the labor force and thus seeking jobs, but who do not have jobs

Patterns of Unemployment

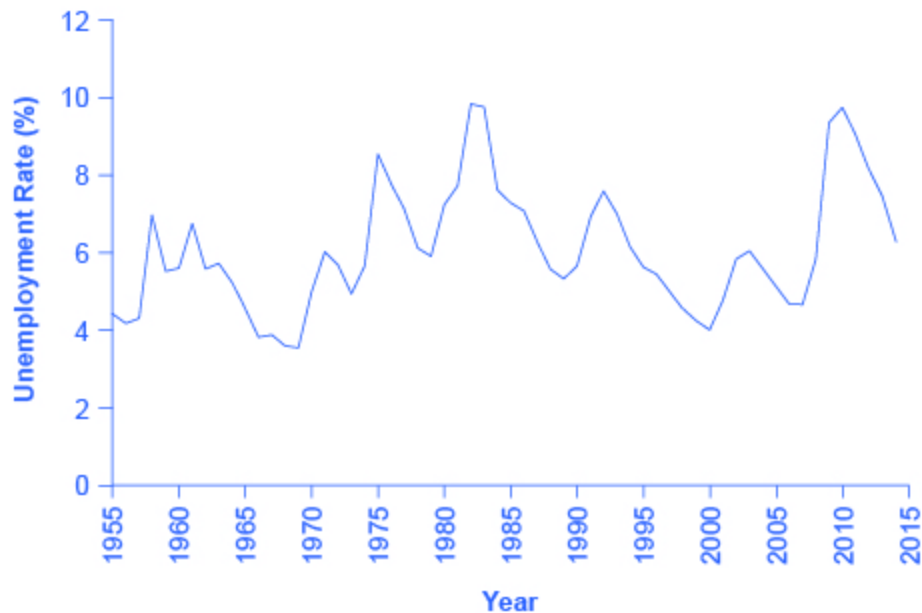
By the end of this section, you will be able to:

- Explain historical patterns of unemployment in the U.S.
- Identify trends of unemployment based on demographics
- Evaluate global unemployment rates

Let's look at how unemployment rates have changed over time and how various groups of people are affected by unemployment differently.

The Historical U.S. Unemployment Rate

[\[link\]](#) shows the historical pattern of U.S. unemployment since 1955. The U.S. Unemployment Rate, 1955–2015



The U.S. unemployment rate moves up and down as the economy moves in and out of recessions. But over time, the unemployment rate seems to return to a range of 4% to 6%. There does not seem to be a long-term trend toward the rate moving generally higher or generally lower. (Source: Federal Reserve Economic Data (FRED) <https://research.stlouisfed.org/fred2/series/LRUN64TTUSA156S0>)

As we look at this data, several patterns stand out:

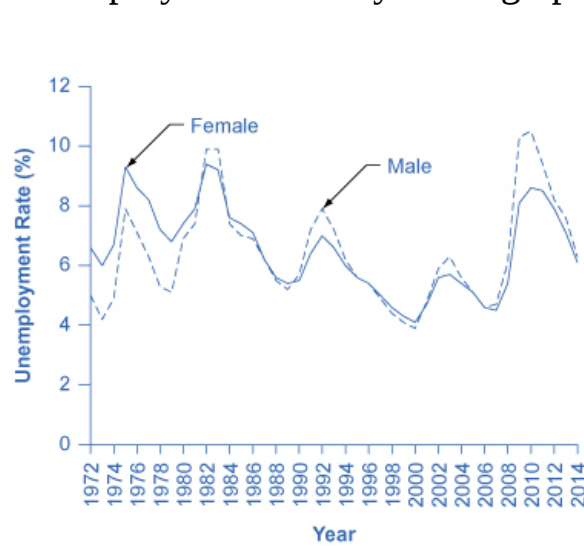
1. Unemployment rates do fluctuate over time. During the deep recessions of the early 1980s and of 2007–2009, unemployment reached roughly 10%. For comparison, during the Great Depression of the 1930s, the unemployment rate reached almost 25% of the labor force.
2. Unemployment rates in the late 1990s and into the mid-2000s were rather low by historical standards. The unemployment rate was below 5% from 1997 to 2000 and near 5% during almost all of 2006–2007. The previous time unemployment had been less than 5% for three consecutive years was three decades earlier, from 1968 to 1970.
3. The unemployment rate never falls all the way to zero. Indeed, it never seems to get below 3%—and it stays that low only for very short periods. (Reasons why this is the case are discussed later in this chapter.)
4. The timing of rises and falls in unemployment matches fairly well with the timing of upswings and downswings in the overall economy. During periods of recession and depression, unemployment is high. During periods of economic growth, unemployment tends to be lower.
5. No significant upward or downward trend in unemployment rates is apparent. This point is especially worth noting because the U.S. population nearly quadrupled from 76 million in 1900 to over 314 million by 2012. Moreover, a higher proportion of U.S. adults are now in the paid workforce, because women have entered the paid labor force in significant numbers in recent decades. Women composed 18% of the paid workforce in 1900 and nearly half of the paid workforce in 2012. But despite the increased number of workers, as well as other economic events like globalization and the continuous invention of new technologies, the economy has provided jobs without causing any long-term upward or downward trend in unemployment rates.

Unemployment Rates by Group

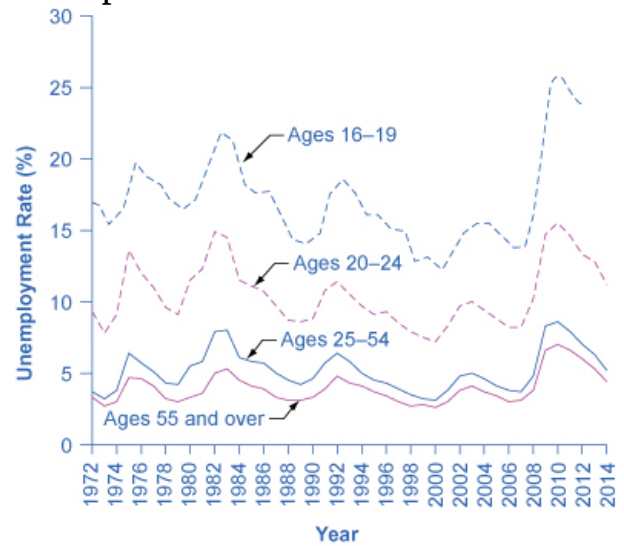
Unemployment is not distributed evenly across the U.S. population. [\[link\]](#) shows unemployment rates broken down in various ways: by gender, age,

and race/ethnicity.

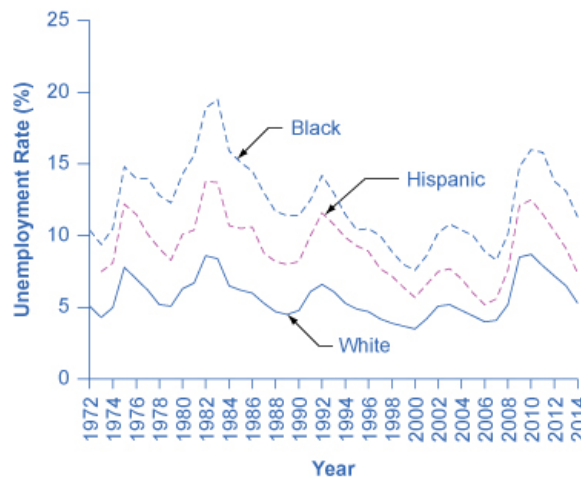
Unemployment Rate by Demographic Group



(a) Unemployment rates by gender



(b) Unemployment rates for women, by age



(c) Unemployment rates by race and ethnicity

(a) By gender, 1972–2014. Unemployment rates for men used to be lower than unemployment rates for women, but in recent decades, the two rates have been very close, often with the unemployment rate for men somewhat higher. (b) By age, 1972–2014. Unemployment rates are highest for the very young and become lower with age. (c) By race and ethnicity, 1972–2014. Although unemployment rates for all groups tend to rise and fall together, the unemployment rate for whites has

been lower than the unemployment rate for blacks and Hispanics in recent decades. (Source: www.bls.gov)

The unemployment rate for women had historically tended to be higher than the unemployment rate for men, perhaps reflecting the historical pattern that women were seen as “secondary” earners. By about 1980, however, the unemployment rate for women was essentially the same as that for men, as shown in [\[link\]](#) (a). During the recession of 2008-2009, the unemployment rate for men exceeded the unemployment rate for women. Through 2014, this pattern has remained, although the gap is narrowing.

Note:

Read this [report](#) for detailed information on the recession of 2008–2009. It also provides some very useful information on the statistics of unemployment.



Younger workers tend to have higher unemployment, while middle-aged workers tend to have lower unemployment, probably because the middle-aged workers feel the responsibility of needing to have a job more heavily. Younger workers move in and out of jobs (and in and out of the labor force) more easily. Elderly workers have extremely low rates of unemployment, because those who do not have jobs often exit the labor force by retiring, and thus are not counted in the unemployment statistics. [\[link\]](#) (b) shows unemployment rates for women divided by age; the pattern for men is similar.

The unemployment rate for African-Americans is substantially higher than the rate for other racial or ethnic groups, a fact that surely reflects, to some extent, a pattern of discrimination that has constrained blacks’ labor market opportunities. However, the gaps between unemployment rates for whites and for blacks and Hispanics diminished in the 1990s, as shown in [\[link\]](#) (c). In fact, unemployment rates for blacks and Hispanics were at the lowest levels for several decades in the mid-2000s before rising during the recent Great Recession.

Finally, those with less education typically suffer higher unemployment. In February 2015, for example, the unemployment rate for those with a college degree was 2.7%; for those with some college but not a four year degree, the unemployment rate was 5.1%; for high school graduates with no additional degree, the unemployment rate was 5.4%; and for those without a high school diploma, the unemployment rate was 8.4%. This pattern may arise because additional education offers better connections to the labor market and higher demand, or it may occur because the labor market opportunities for low-skilled workers are less attractive than the opportunities for the more highly-skilled. Because of lower pay, low-skilled workers may be less motivated to find jobs.

Breaking Down Unemployment in Other Ways

The Bureau of Labor Statistics also gives information about the reasons for being unemployed as well as the length of time individuals have been unemployed. [\[link\]](#), for example, shows the four reasons for being unemployed and the percentages of the currently unemployed that fall into each category. [\[link\]](#) shows the length of unemployment. For both of these, the data is from February of 2015. (bls.gov)

Reason	Percentage

Reason	Percentage
New Entrants	11.2%
Re-entrants	30.5%
Job Leavers	10.2%
Job Losers: Temporary	11.7%
Job Losers: Non Temporary	36.3%

Reasons for Being Unemployed, February 2015

Length of Time	Percentage
Under 5 weeks	27.9%
5 to 14 weeks	25.6%
15 to 26 weeks	15.4%
Over 27 weeks	31.1%

Length of Unemployment, February 2015

Note:

Watch this [speech](#) on the impact of droids on the labor market.



International Unemployment Comparisons

From an international perspective, the U.S. unemployment rate typically has looked a little better than average. [\[link\]](#) compares unemployment rates for 1991, 1996, 2001, 2006 (just before the recession), and 2012 (somewhat after the recession) from several other high-income countries.

Country	1991	1996	2001	2006	2012
United States	6.8%	5.4%	4.8%	4.4%	8.1%
Canada	9.8%	8.8%	6.4%	6.2%	6.3%
Japan	2.1%	3.4%	5.1%	4.5%	3.9%
France	9.5%	12.5%	8.7%	10.1%	10.0%
Germany	5.6%	9.0%	8.9%	9.8%	5.5%
Italy	6.9%	11.7%	9.6%	7.8%	10.8%
Sweden	3.1%	9.9%	5.0%	5.2%	7.9%

Country	1991	1996	2001	2006	2012
United Kingdom	8.8%	8.1%	5.1%	5.5%	8.0%

International Comparisons of Unemployment Rates

However, cross-country comparisons of unemployment rates need to be treated with care, because each country has slightly different survey tools for measuring unemployment and also different labor markets. For example, Japan's unemployment rates appear quite low, but Japan's economy has been mired in slow growth and recession since the late 1980s, and Japan's unemployment rate probably paints too rosy a picture of its labor market. In Japan, workers who lose their jobs are often quick to exit the labor force and not look for a new job, in which case they are not counted as unemployed. In addition, Japanese firms are often quite reluctant to fire workers, and so firms have substantial numbers of workers who are on reduced hours or officially employed, but doing very little. This Japanese pattern is perhaps best viewed as an unusual method for society to provide support for the unemployed, rather than a sign of a healthy economy.

Note:

We hear about the Chinese economy in the news all the time. The value of the Chinese yuan in comparison to the U.S. dollar is likely to be part of the nightly business report. So why is the Chinese economy not included in this discussion of international unemployment? The lack of reliable statistics is probably the reason. This [article](#) explains why.



Comparing unemployment rates in the United States and other high-income economies with unemployment rates in Latin America, Africa, Eastern Europe, and Asia is very difficult. One reason is that the statistical agencies in many poorer countries lack the resources and technical capabilities of the U.S. Bureau of the Census. But a more difficult problem with international comparisons is that in many low-income countries, most workers are not involved in the labor market through an employer who pays them regularly. Instead, workers in these countries are engaged in short-term work, subsistence activities, and barter. Moreover, the effect of unemployment is very different in high-income and low-income countries. Unemployed workers in the developed economies have access to various government programs like unemployment insurance, welfare, and food stamps; such programs may barely exist in poorer countries. Although unemployment is a serious problem in many low-income countries, it manifests itself in a different way than in high-income countries.

Key Concepts and Summary

The U.S. unemployment rate rises during periods of recession and depression, but falls back to the range of 4% to 6% when the economy is strong. The unemployment rate never falls to zero. Despite enormous growth in the size of the U.S. population and labor force in the twentieth century, along with other major trends like globalization and new technology, the unemployment rate shows no long-term rising trend.

Unemployment rates differ by group: higher for African-Americans and Hispanics than for whites; higher for less educated than more educated; higher for the young than the middle-aged. Women's unemployment rates used to be higher than men's, but in recent years men's and women's unemployment rates have been very similar. In recent years, unemployment rates in the United States have compared favorably with unemployment rates in most other high-income economies.

Self-Check Questions

Exercise:**Problem:**

Over the long term, has the U.S. unemployment rate generally trended up, trended down, or remained at basically the same level?

Solution:

Over the long term, the U.S. unemployment rate has remained basically the same level.

Exercise:**Problem:**

Whose unemployment rates are commonly higher in the U.S. economy:

- a. Whites or nonwhites?
 - b. The young or the middle-aged?
 - c. College graduates or high school graduates?
-

Solution:

- a. Nonwhites
- b. The young
- c. High school graduates

Review Questions**Exercise:****Problem:**

Are U.S. unemployment rates typically higher, lower, or about the same as unemployment rates in other high-income countries?

Exercise:**Problem:**

Are U.S. unemployment rates distributed evenly across the population?

Critical Thinking Questions**Exercise:****Problem:**

Is the higher unemployment rates for minority workers necessarily an indication of discrimination? What could be some other reasons for the higher unemployment rate?

Exercise:**Problem:**

While unemployment is highly negatively correlated with the level of economic activity, in the real world it responds with a lag. In other words, firms do not immediately lay off workers in response to a sales decline. They wait a while before responding. Similarly, firms do not immediately hire workers when sales pick up. What do you think accounts for the lag in response time?

Exercise:**Problem:**

Why do you think that unemployment rates are lower for individuals with more education?

What Causes Changes in Unemployment over the Short Run

By the end of this section, you will be able to:

- Analyze cyclical unemployment
- Explain the relationship between sticky wages and employment using various economic arguments
- Apply supply and demand models to unemployment and wages

We have seen that unemployment varies across times and places. What causes changes in unemployment? There are different answers in the short run and in the long run. Let's look at the short run first.

Cyclical Unemployment

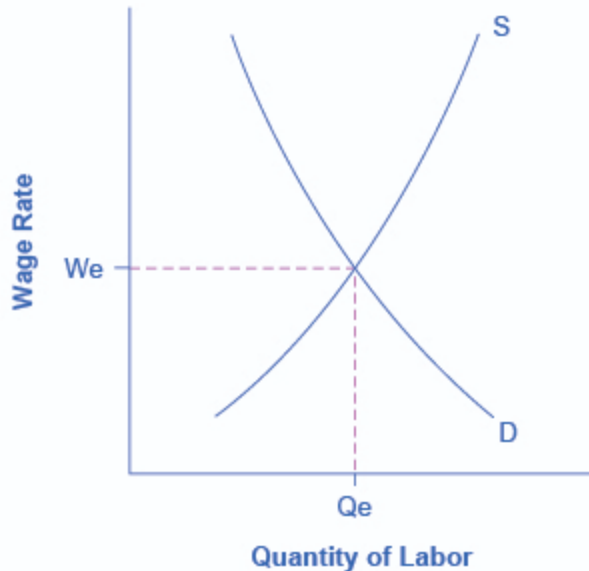
Let's make the plausible assumption that in the short run, from a few months to a few years, the quantity of hours that the average person is willing to work for a given wage does not change much, so the labor supply curve does not shift much. In addition, make the standard *ceteris paribus* assumption that there is no substantial short-term change in the age structure of the labor force, institutions and laws affecting the labor market, or other possibly relevant factors.

One primary determinant of the demand for labor from firms is how they perceive the state of the macro economy. If firms believe that business is expanding, then at any given wage they will desire to hire a greater quantity of labor, and the labor demand curve shifts to the right. Conversely, if firms perceive that the economy is slowing down or entering a recession, then they will wish to hire a lower quantity of labor at any given wage, and the labor demand curve will shift to the left. The variation in unemployment caused by the economy moving from expansion to recession or from recession to expansion (i.e. the business cycle) is known as **cyclical unemployment**.

From the standpoint of the supply-and-demand model of competitive and flexible labor markets, unemployment represents something of a puzzle. In a supply-and-demand model of a labor market, as illustrated in [\[link\]](#), the

labor market should move toward an equilibrium wage and quantity. At the equilibrium wage (W_e), the equilibrium quantity (Q_e) of labor supplied by workers should be equal to the quantity of labor demanded by employers.

The Unemployment and Equilibrium in the Labor Market



In a labor market with flexible wages, the equilibrium will occur at wage W_e and quantity Q_e , where the number of people looking for jobs (shown by S) equals the number of jobs available (shown by D).

One possibility for unemployment is that people who are unemployed are those who are not willing to work at the current equilibrium wage, say \$10 an hour, but would be willing to work at a higher wage, like \$20 per hour. The monthly Current Population Survey would count these people as unemployed, because they say they are ready and looking for work (at \$20 per hour). But from an economist's point of view, these people are choosing to be unemployed.

Probably a few people are unemployed because of unrealistic expectations about wages, but they do not represent the majority of the unemployed. Instead, unemployed people often have friends or acquaintances of similar

skill levels who are employed, and the unemployed would be willing to work at the jobs and wages similar to what is being received by those people. But the employers of their friends and acquaintances do not seem to be hiring. In other words, these people are involuntarily unemployed. What causes involuntary unemployment?

Why Wages Might Be Sticky Downward

If a labor market model with flexible wages does not describe unemployment very well—because it predicts that anyone willing to work at the going wage can always find a job—then it may prove useful to consider economic models in which wages are not flexible or adjust only very slowly. In particular, even though wage increases may occur with relative ease, wage decreases are few and far between.

One set of reasons why wages may be “sticky downward,” as economists put it, involves economic laws and institutions. For low-skilled workers being paid the minimum wage, it is illegal to reduce their wages. For union workers operating under a multiyear contract with a company, wage cuts might violate the contract and create a labor dispute or a strike. However, minimum wages and union contracts are not a sufficient reason why wages would be sticky downward for the U.S. economy as a whole. After all, out of the 150 million or so workers in the U.S. economy, only about 1.4 million—less than 2% of the total—are paid the minimum wage. Similarly, only about 12% of American wage and salary workers are represented by a labor union. In other high-income countries, more workers may have their wages determined by unions or the minimum wage may be set at a level that applies to a larger share of workers. But for the United States, these two factors combined affect only about one-fifth or less of the labor force.

Economists looking for reasons why wages might be sticky downwards have focused on factors that may characterize most labor relationships in the economy, not just a few. A number of different theories have been proposed, but they share a common tone.

One argument is that even employees who are not union members often work under an **implicit contract**, which is that the employer will try to

keep wages from falling when the economy is weak or the business is having trouble, and the employee will not expect huge salary increases when the economy or the business is strong. This wage-setting behavior acts like a form of insurance: the employee has some protection against wage declines in bad times, but pays for that protection with lower wages in good times. Clearly, this sort of implicit contract means that firms will be hesitant to cut wages, lest workers feel betrayed and work less hard or even leave the firm.

Efficiency wage theory argues that the productivity of workers depends on their pay, and so employers will often find it worthwhile to pay their employees somewhat more than market conditions might dictate. One reason is that employees who are paid better than others will be more productive because they recognize that if they were to lose their current jobs, they would suffer a decline in salary. As a result, they are motivated to work harder and to stay with the current employer. In addition, employers know that it is costly and time-consuming to hire and train new employees, so they would prefer to pay workers a little extra now rather than to lose them and have to hire and train new workers. Thus, by avoiding wage cuts, the employer minimizes costs of training and hiring new workers, and reaps the benefits of well-motivated employees.

The **adverse selection of wage cuts argument** points out that if an employer reacts to poor business conditions by reducing wages for all workers, then the best workers, those with the best employment alternatives at other firms, are the most likely to leave. The least attractive workers, with fewer employment alternatives, are more likely to stay. Consequently, firms are more likely to choose which workers should depart, through layoffs and firings, rather than trimming wages across the board. Sometimes companies that are going through tough times can persuade workers to take a pay cut for the short term, and still retain most of the firm's workers. But these stories are notable because they are so uncommon. It is far more typical for companies to lay off some workers, rather than to cut wages for everyone.

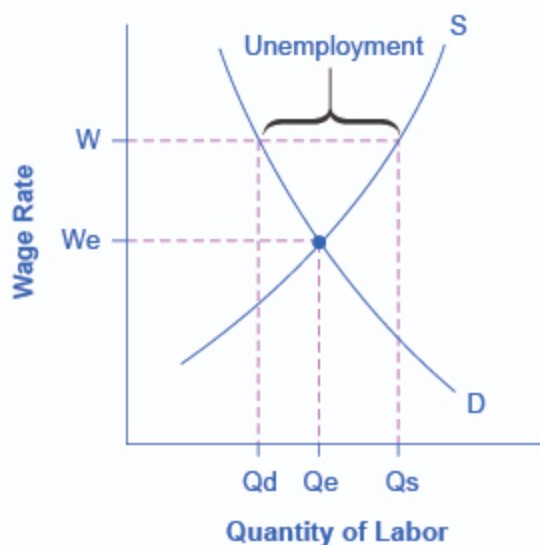
The **insider-outsider model** of the labor force, in simple terms, argues that those already working for firms are "insiders," while new employees, at

least for a time, are “outsiders.” A firm depends on its insiders to grease the wheels of the organization, to be familiar with routine procedures, to train new employees, and so on. However, cutting wages will alienate the insiders and damage the firm’s productivity and prospects.

Finally, the **relative wage coordination argument** points out that even if most workers were hypothetically willing to see a decline in their own wages in bad economic times as long as everyone else also experiences such a decline, there is no obvious way for a decentralized economy to implement such a plan. Instead, workers confronted with the possibility of a wage cut will worry that other workers will not have such a wage cut, and so a wage cut means being worse off both in absolute terms and relative to others. As a result, workers fight hard against wage cuts.

These theories of why wages tend not to move downward differ in their logic and their implications, and figuring out the strengths and weaknesses of each theory is an ongoing subject of research and controversy among economists. All tend to imply that wages will decline only very slowly, if at all, even when the economy or a business is having tough times. When wages are inflexible and unlikely to fall, then either short-run or long-run unemployment can result. This can be seen in [\[link\]](#).

Sticky Wages in the Labor Market



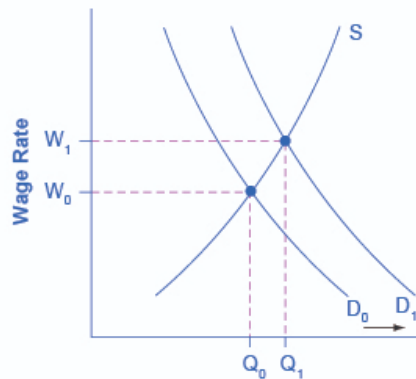
Because the wage rate is stuck at W , above

the equilibrium, the number of job seekers (Q_s) is greater than the number of job openings (Q_d). The result is unemployment, shown by the bracket in the figure.

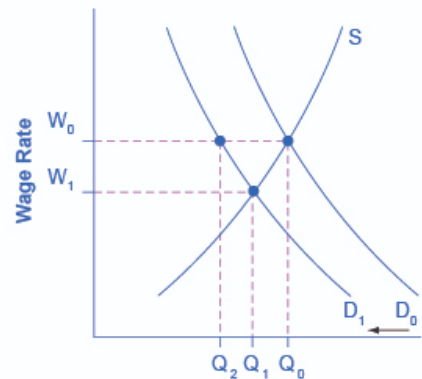
The interaction between shifts in labor demand and wages that are sticky downward are shown in [\[link\]](#). [\[link\]](#) (a) illustrates the situation in which the demand for labor shifts to the right from D_0 to D_1 . In this case, the equilibrium wage rises from W_0 to W_1 and the equilibrium quantity of labor hired increases from Q_0 to Q_1 . It does not hurt employee morale at all for wages to rise.

[\[link\]](#) (b) shows the situation in which the demand for labor shifts to the left, from D_0 to D_1 , as it would tend to do in a recession. Because wages are sticky downward, they do not adjust toward what would have been the new equilibrium wage (W_1), at least not in the short run. Instead, after the shift in the labor demand curve, the same quantity of workers is willing to work at that wage as before; however, the quantity of workers demanded at that wage has declined from the original equilibrium (Q_0) to Q_2 . The gap between the original equilibrium quantity (Q_0) and the new quantity demanded of labor (Q_2) represents workers who would be willing to work at the going wage but cannot find jobs. The gap represents the economic meaning of unemployment.

Rising Wage and Low Unemployment: Where Is the Unemployment in Supply and Demand?



(a) Rising demand for labor, wages rise



(b) Falling demand for labor, sticky wages, and unemployment

(a) In a labor market where wages are able to rise, an increase in the demand for labor from D_0 to D_1 leads to an increase in equilibrium quantity of labor hired from Q_0 to Q_1 and a rise in the equilibrium wage from W_0 to W_1 . (b) In a labor market where wages do not decline, a fall in the demand for labor from D_0 to D_1 leads to a decline in the quantity of labor demanded at the original wage (W_0) from Q_0 to Q_2 . These workers will want to work at the prevailing wage (W_0), but will not be able to find jobs.

This analysis helps to explain the connection noted earlier: that unemployment tends to rise in recessions and to decline during expansions. The overall state of the economy shifts the labor demand curve and, combined with wages that are sticky downwards, unemployment changes. The rise in unemployment that occurs because of a recession is cyclical unemployment.

Note:

The St. Louis Federal Reserve Bank is the best resource for macroeconomic time series data, known as the Federal Reserve Economic Data (FRED). [FRED](https://fred.stlouisfed.org/) provides complete data sets on various measures of

the unemployment rate as well as the monthly Bureau of Labor Statistics report on the results of the household and employment surveys.



Key Concepts and Summary

Cyclical unemployment rises and falls with the business cycle. In a labor market with flexible wages, wages will adjust in such a market so that quantity demanded of labor always equals the quantity supplied of labor at the equilibrium wage. Many theories have been proposed for why wages might not be flexible, but instead may adjust only in a “sticky” way, especially when it comes to downward adjustments: implicit contracts, efficiency wage theory, adverse selection of wage cuts, insider-outsider model, and relative wage coordination.

Self-Check Questions

Exercise:

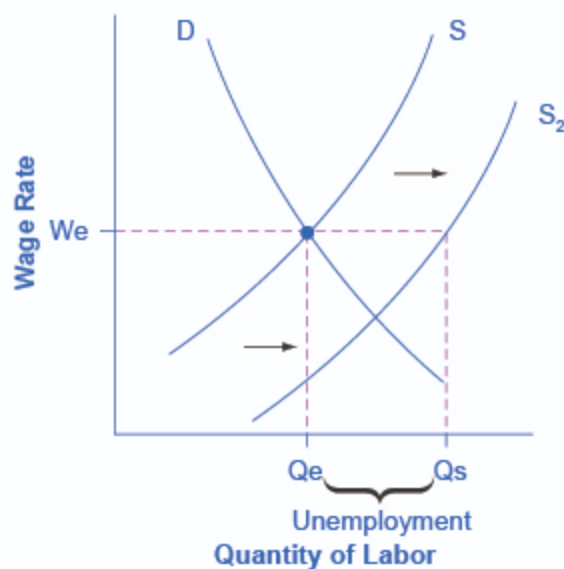
Problem:

Beginning in the 1970s and continuing for three decades, women entered the U.S. labor force in a big way. If we assume that wages are sticky in a downward direction, but that around 1970 the demand for labor equaled the supply of labor at the current wage rate, what do you imagine happened to the wage rate, employment, and unemployment as a result of increased labor force participation?

Solution:

Because of the influx of women into the labor market, the supply of labor shifts to the right. Since wages are sticky downward, the increased supply of labor causes an increase in people looking for jobs (Q_s), but no change in the number of jobs available (Q_e). As a result, unemployment increases by the amount of the increase in the labor supply. This can be seen in the following figure.

Over time, as labor demand grows, the unemployment will decline and eventually wages will begin to increase again. But this increase in labor demand goes beyond the scope of this problem.



Review Questions

Exercise:

Problem:

When would you expect cyclical unemployment to be rising? Falling?

Exercise:

Problem:

Why is there unemployment in a labor market with flexible wages?

Exercise:**Problem:**

Name and explain some of the reasons why wages are likely to be sticky, especially in downward adjustments.

Critical Thinking Questions**Exercise:****Problem:**

Do you think it is rational for workers to prefer sticky wages to wage cuts, when the consequence of sticky wages is unemployment for some workers? Why or why not? How do the reasons for sticky wages explained in this section apply to your argument?

Problems**Exercise:****Problem:**

A government passes a family-friendly law that no companies can have evening, nighttime, or weekend hours, so that everyone can be home with their families during these times. Analyze the effect of this law using a demand and supply diagram for the labor market: first assuming that wages are flexible, and then assuming that wages are sticky downward.

Glossary

adverse selection of wage cuts argument

if employers reduce wages for all workers, the best will leave

cyclical unemployment

unemployment closely tied to the business cycle, like higher unemployment during a recession

efficiency wage theory

the theory that the productivity of workers, either individually or as a group, will increase if they are paid more

implicit contract

an unwritten agreement in the labor market that the employer will try to keep wages from falling when the economy is weak or the business is having trouble, and the employee will not expect huge salary increases when the economy or the business is strong

insider-outsider model

those already working for the firm are “insiders” who know the procedures; the other workers are “outsiders” who are recent or prospective hires

relative wage coordination argument

across-the-board wage cuts are hard for an economy to implement, and workers fight against them

What Causes Changes in Unemployment over the Long Run

By the end of this section, you will be able to:

- Explain frictional and structural unemployment
- Assess relationships between the natural rate of employment and potential real GDP, productivity, and public policy
- Identify recent patterns in the natural rate of employment
- Propose ways to combat unemployment

Cyclical unemployment explains why unemployment rises during a recession and falls during an economic expansion. But what explains the remaining level of unemployment even in good economic times? Why is the unemployment rate never zero? Even when the U.S. economy is growing strongly, the unemployment rate only rarely dips as low as 4%. Moreover, the discussion earlier in this chapter pointed out that unemployment rates in many European countries like Italy, France, and Germany have often been remarkably high at various times in the last few decades. Why does some level of unemployment persist even when economies are growing strongly? Why are unemployment rates continually higher in certain economies, through good economic years and bad? Economists have a term to describe the remaining level of unemployment that occurs even when the economy is healthy: it is called the **natural rate of unemployment**.

The Long Run: The Natural Rate of Unemployment

The natural rate of unemployment is not “natural” in the sense that water freezes at 32 degrees Fahrenheit or boils at 212 degrees Fahrenheit. It is not a physical and unchanging law of nature. Instead, it is only the “natural” rate because it is the unemployment rate that would result from the combination of economic, social, and political factors that exist at a time—assuming the economy was neither booming nor in recession. These forces include the usual pattern of companies expanding and contracting their workforces in a dynamic economy, social and economic forces that affect the labor market, or public policies that affect either the eagerness of people

to work or the willingness of businesses to hire. Let's discuss these factors in more detail.

Frictional Unemployment

In a market economy, some companies are always going broke for a variety of reasons: old technology; poor management; good management that happened to make bad decisions; shifts in tastes of consumers so that less of the firm's product is desired; a large customer who went broke; or tough domestic or foreign competitors. Conversely, other companies will be doing very well for just the opposite reasons and looking to hire more employees. In a perfect world, all of those who lost jobs would immediately find new ones. But in the real world, even if the number of job seekers is equal to the number of job vacancies, it takes time to find out about new jobs, to interview and figure out if the new job is a good match, or perhaps to sell a house and buy another in proximity to a new job. The unemployment that occurs in the meantime, as workers move between jobs, is called **frictional unemployment**. Frictional unemployment is not inherently a bad thing. It takes time on part of both the employer and the individual to match those looking for employment with the correct job openings. For individuals and companies to be successful and productive, you want people to find the job for which they are best suited, not just the first job offered.

In the mid-2000s, before the recession of 2008–2009, it was true that about 7% of U.S. workers saw their jobs disappear in any three-month period. But in periods of economic growth, these destroyed jobs are counterbalanced for the economy as a whole by a larger number of jobs created. In 2005, for example, there were typically about 7.5 million unemployed people at any given time in the U.S. economy. Even though about two-thirds of those unemployed people found a job in 14 weeks or fewer, the unemployment rate did not change much during the year, because those who found new jobs were largely offset by others who lost jobs.

Of course, it would be preferable if people who were losing jobs could immediately and easily move into the new jobs being created, but in the real world, that is not possible. Someone who is laid off by a textile mill in South Carolina cannot turn around and immediately start working for a

textile mill in California. Instead, the adjustment process happens in ripples. Some people find new jobs near their old ones, while others find that they must move to new locations. Some people can do a very similar job with a different company, while others must start new career paths. Some people may be near retirement and decide to look only for part-time work, while others want an employer that offers a long-term career path. The frictional unemployment that results from people moving between jobs in a dynamic economy may account for one to two percentage points of total unemployment.

The level of frictional unemployment will depend on how easy it is for workers to learn about alternative jobs, which may reflect the ease of communications about job prospects in the economy. The extent of frictional unemployment will also depend to some extent on how willing people are to move to new areas to find jobs—which in turn may depend on history and culture.

Frictional unemployment and the natural rate of unemployment also seem to depend on the age distribution of the population. [\[link\]](#) (b) showed that unemployment rates are typically lower for people between 25–54 years of age than they are for those who are either younger or older. “Prime-age workers,” as those in the 25–54 age bracket are sometimes called, are typically at a place in their lives when they want to have a job and income arriving at all times. But some proportion of those who are under 30 may still be trying out jobs and life options and some proportion of those over 55 are eyeing retirement. In both cases, the relatively young or old tend to worry less about unemployment than those in-between, and their periods of frictional unemployment may be longer as a result. Thus, a society with a relatively high proportion of relatively young or old workers will tend to have a higher unemployment rate than a society with a higher proportion of its workers in middle age.

Structural Unemployment

Another factor that influences the natural rate of unemployment is the amount of **structural unemployment**. The structurally unemployed are individuals who have no jobs because they lack skills valued by the labor

market, either because demand has shifted away from the skills they do have, or because they never learned any skills. An example of the former would be the unemployment among aerospace engineers after the U.S. space program downsized in the 1970s. An example of the latter would be high school dropouts.

Some people worry that technology causes structural unemployment. In the past, new technologies have put lower skilled employees out of work, but at the same time they create demand for higher skilled workers to use the new technologies. Education seems to be the key in minimizing the amount of structural unemployment. Individuals who have degrees can be retrained if they become structurally unemployed. For people with no skills and little education, that option is more limited.

Natural Unemployment and Potential Real GDP

The natural unemployment rate is related to two other important concepts: full employment and potential real GDP. The economy is considered to be at full employment when the actual unemployment rate is equal to the natural unemployment. When the economy is at full employment, real GDP is equal to potential real GDP. By contrast, when the economy is below full employment, the unemployment rate is greater than the natural unemployment rate and real GDP is less than potential. Finally, when the economy is above full employment, then the unemployment rate is less than the natural unemployment rate and real GDP is greater than potential. Operating above potential is only possible for a short while, since it is analogous to all workers working overtime.

Productivity Shifts and the Natural Rate of Unemployment

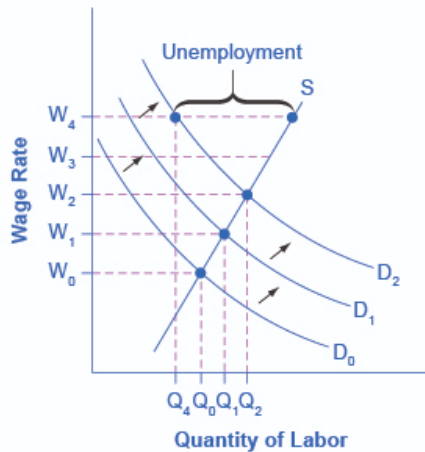
Unexpected shifts in productivity can have a powerful effect on the natural rate of unemployment. Over time, the level of wages in an economy will be determined by the productivity of workers. After all, if a business paid workers more than could be justified by their productivity, the business will ultimately lose money and go bankrupt. Conversely, if a business tries to pay workers less than their productivity then, in a competitive labor market,

other businesses will find it worthwhile to hire away those workers and pay them more.

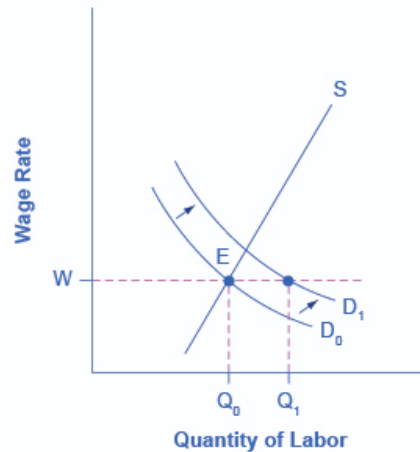
However, adjustments of wages to productivity levels will not happen quickly or smoothly. Wages are typically reviewed only once or twice a year. In many modern jobs, it is difficult to measure productivity at the individual level. For example, how precisely would one measure the quantity produced by an accountant who is one of many people working in the tax department of a large corporation? Because productivity is difficult to observe, wage increases are often determined based on recent experience with productivity; if productivity has been rising at, say, 2% per year, then wages rise at that level as well. However, when productivity changes unexpectedly, it can affect the natural rate of unemployment for a time.

The U.S. economy in the 1970s and 1990s provides two vivid examples of this process. In the 1970s, productivity growth slowed down unexpectedly (as discussed in [Economic Growth](#)). For example, output per hour of U.S. workers in the business sector increased at an annual rate of 3.3% per year from 1960 to 1973, but only 0.8% from 1973 to 1982. [\[link\]](#) (a) illustrates the situation where the demand for labor—that is, the quantity of labor that business is willing to hire at any given wage—has been shifting out a little each year because of rising productivity, from D_0 to D_1 to D_2 . As a result, equilibrium wages have been rising each year from W_0 to W_1 to W_2 . But when productivity unexpectedly slows down, the pattern of wage increases does not adjust right away. Wages keep rising each year from W_2 to W_3 to W_4 . But the demand for labor is no longer shifting up. A gap opens where the quantity of labor supplied at wage level W_4 is greater than the quantity demanded. The natural rate of unemployment rises; indeed, in the aftermath of this unexpectedly low productivity in the 1970s, the national unemployment rate did not fall below 7% from May, 1980 until 1986. Over time, the rise in wages will adjust to match the slower gains in productivity, and the unemployment rate will ease back down. But this process may take years.

Unexpected Productivity Changes and Unemployment



(a) Productivity rises, and then stops rising



(b) Productivity doesn't change, and then rises

(a) Productivity is rising, increasing the demand for labor. Employers and workers become used to the pattern of wage increases. Then productivity suddenly stops increasing. However, the expectations of employers and workers for wage increases do not shift immediately, so wages keep rising as before. But the demand for labor has not increased, so at wage W_4 , unemployment exists where the quantity supplied of labor exceeds the quantity demanded. (b) The rate of productivity increase has been zero for a time, so employers and workers have come to accept the equilibrium wage level (W). Then productivity increases unexpectedly, shifting demand for labor from D_0 to D_1 . At the wage (W), this means that the quantity demanded of labor exceeds the quantity supplied, and with job offers plentiful, the unemployment rate will be low.

The late 1990s provide an opposite example: instead of the surprise decline in productivity in the 1970s, productivity unexpectedly rose in the mid-1990s. The annual growth rate of real output per hour of labor increased from 1.7% from 1980–1995, to an annual rate of 2.6% from 1995–2001. Let's simplify the situation a bit, so that the economic lesson of the story is easier to see graphically, and say that productivity had not been increasing at all in earlier years, so the intersection of the labor market was at point E in [link](#) (b), where the demand curve for labor (D_0) intersects the supply curve for labor. As a result, real wages were not increasing. Now,

productivity jumps upward, which shifts the demand for labor out to the right, from D_0 to D_1 . At least for a time, however, wages are still being set according to the earlier expectations of no productivity growth, so wages do not rise. The result is that at the prevailing wage level (W), the quantity of labor demanded (Q_d) will for a time exceed the quantity of labor supplied (Q_s), and unemployment will be very low—actually below the natural level of unemployment for a time. This pattern of unexpectedly high productivity helps to explain why the unemployment rate stayed below 4.5%—quite a low level by historical standards—from 1998 until after the U.S. economy had entered a recession in 2001.

Average levels of unemployment will tend to be somewhat higher on average when productivity is unexpectedly low, and conversely, will tend to be somewhat lower on average when productivity is unexpectedly high. But over time, wages do eventually adjust to reflect productivity levels.

Public Policy and the Natural Rate of Unemployment

Public policy can also have a powerful effect on the natural rate of unemployment. On the supply side of the labor market, public policies to assist the unemployed can affect how eager people are to find work. For example, if a worker who loses a job is guaranteed a generous package of unemployment insurance, welfare benefits, food stamps, and government medical benefits, then the opportunity cost of being unemployed is lower and that worker will be less eager to seek a new job.

What seems to matter most is not just the amount of these benefits, but how long they last. A society that provides generous help for the unemployed that cuts off after, say, six months, may provide less of an incentive for unemployment than a society that provides less generous help that lasts for several years. Conversely, government assistance for job search or retraining can in some cases encourage people back to work sooner. See the Clear it Up to learn how the U.S. handles unemployment insurance.

Note:

How does U.S. unemployment insurance work?

Unemployment insurance is a joint federal–state program, established by federal law in 1935. The federal government sets minimum standards for the program, but most of the administration is done by state governments. The funding for the program is a federal tax collected from employers. The federal government requires that the tax be collected on the first \$7,000 in wages paid to each worker; however, states can choose to collect the tax on a higher amount if they wish, and 41 states have set a higher limit. States can choose the length of time that benefits will be paid, although most states limit unemployment benefits to 26 weeks—with extensions possible in times of especially high unemployment. The fund is then used to pay benefits to those who become unemployed. Average unemployment benefits are equal to about one-third of the wage earned by the person in his or her previous job, but the level of unemployment benefits varies considerably across states.

Bottom 10 States That Pay the Lowest Benefit per Week		Top 10 States That Pay the Highest Benefit per Week	
Delaware	\$330	Massachusetts	\$674
Georgia	\$330	Minnesota	\$629
South Carolina	\$326	New Jersey	\$624
Missouri	\$320	Washington	\$624
Florida	\$275	Connecticut	\$590
Tennessee	\$275	Pennsylvania	\$573
Alabama	\$265	Rhode Island	\$566

Bottom 10 States That Pay the Lowest Benefit per Week		Top 10 States That Pay the Highest Benefit per Week	
Louisiana	\$247	Ohio	\$564
Arizona	\$240	Hawaii	\$560
Mississippi	\$235	Oregon	\$538

Maximum Weekly Unemployment Benefits by State in 2014(Source: <http://jobsearch.about.com/od/unemployment/fl/unemployment-benefits-by-state-2014.htm>)

One other interesting thing to note about the classifications of unemployment—an individual does not have to collect unemployment benefits to be classified as unemployed. While there are statistics kept and studied relating to how many people are collecting unemployment insurance, this is not the source of unemployment rate information.

Note:

View this [article](#) for an explanation of exactly who is eligible for unemployment benefits.



On the demand side of the labor market, government rules social institutions, and the presence of unions can affect the willingness of firms to hire. For example, if a government makes it hard for businesses to start up

or to expand, by wrapping new businesses in bureaucratic red tape, then businesses will become more discouraged about hiring. Government regulations can make it harder to start a business by requiring that a new business obtain many permits and pay many fees, or by restricting the types and quality of products that can be sold. Other government regulations, like zoning laws, may limit where business can be done, or whether businesses are allowed to be open during evenings or on Sunday.

Whatever defenses may be offered for such laws in terms of social value—like the value some Christians place on not working on Sunday—these kinds of restrictions impose a barrier between some willing workers and other willing employers, and thus contribute to a higher natural rate of unemployment. Similarly, if government makes it difficult to fire or lay off workers, businesses may react by trying not to hire more workers than strictly necessary—since laying these workers off would be costly and difficult. High minimum wages may discourage businesses from hiring low-skill workers. Government rules may encourage and support powerful unions, which can then push up wages for union workers, but at a cost of discouraging businesses from hiring those workers.

The Natural Rate of Unemployment in Recent Years

The underlying economic, social, and political factors that determine the natural rate of unemployment can change over time, which means that the natural rate of unemployment can change over time, too.

Estimates by economists of the natural rate of unemployment in the U.S. economy in the early 2000s run at about 4.5 to 5.5%. This is a lower estimate than earlier. Three of the common reasons proposed by economists for this change are outlined below.

1. The Internet has provided a remarkable new tool through which job seekers can find out about jobs at different companies and can make contact with relative ease. An Internet search is far easier than trying to find a list of local employers and then hunting up phone numbers for all of their human resources departments, requesting a list of jobs and

application forms, and so on. Social networking sites such as LinkedIn have changed how people find work as well.

2. The growth of the temporary worker industry has probably helped to reduce the natural rate of unemployment. In the early 1980s, only about 0.5% of all workers held jobs through temp agencies; by the early 2000s, the figure had risen above 2%. Temp agencies can provide jobs for workers while they are looking for permanent work. They can also serve as a clearinghouse, helping workers find out about jobs with certain employers and getting a tryout with the employer. For many workers, a temp job is a stepping-stone to a permanent job that they might not have heard about or gotten any other way, so the growth of temp jobs will also tend to reduce frictional unemployment.
3. The aging of the “baby boom generation”—the especially large generation of Americans born between 1946 and 1963—meant that the proportion of young workers in the economy was relatively high in the 1970s, as the boomers entered the labor market, but is relatively low today. As noted earlier, middle-aged workers are far more likely to keep steady jobs than younger workers, a factor that tends to reduce the natural rate of unemployment.

The combined result of these factors is that the natural rate of unemployment was on average lower in the 1990s and the early 2000s than in the 1980s. The Great Recession of 2008–2009 pushed monthly unemployment rates above 10% in late 2009. But even at that time, the Congressional Budget Office was forecasting that by 2015, unemployment rates would fall back to about 5%—lower than it currently is, though not by much. As of early 2015, policymakers still think that unemployment has not yet reached its natural rate.

The Natural Rate of Unemployment in Europe

By the standards of other high-income economies, the natural rate of unemployment in the U.S. economy appears relatively low. Through good economic years and bad, many European economies have had unemployment rates hovering near 10%, or even higher, since the 1970s. European rates of unemployment have been higher not because recessions in Europe have been deeper, but rather because the conditions underlying

supply and demand for labor have been different in Europe, in a way that has created a much higher natural rate of unemployment.

Many European countries have a combination of generous welfare and unemployment benefits, together with nests of rules that impose additional costs on businesses when they hire. In addition, many countries have laws that require firms to give workers months of notice before laying them off and to provide substantial severance or retraining packages after laying them off. The legally required notice before laying off a worker can be more than three months in Spain, Germany, Denmark, and Belgium, and the legally required severance package can be as high as a year's salary or more in Austria, Spain, Portugal, Italy, and Greece. Such laws will surely discourage laying off or firing current workers. But when companies know that it will be difficult to fire or lay off workers, they also become hesitant about hiring in the first place.

The typically higher levels of unemployment in many European countries in recent years, which have prevailed even when economies are growing at a solid pace, are attributable to the fact that the sorts of laws and regulations that lead to a high natural rate of unemployment are much more prevalent in Europe than in the United States.

A Preview of Policies to Fight Unemployment

The [Government Budgets and Fiscal Policy](#) and [Macroeconomic Policy Around the World](#) chapters provide a detailed discussion of how to fight unemployment, when these policies can be discussed in the context of the full array of macroeconomic goals and frameworks for analysis. But even at this preliminary stage, it is useful to preview the main issues concerning policies to fight unemployment.

The remedy for unemployment will depend on the diagnosis. Cyclical unemployment is a short-term problem, caused because the economy is in a recession. Thus, the preferred solution will be to avoid or minimize recessions. As [Government Budgets and Fiscal Policy](#) discusses, this policy can be enacted by stimulating the overall buying power in the economy, so

that firms perceive that sales and profits are possible, which makes them eager to hire.

Dealing with the natural rate of unemployment is trickier. There is not much to be done about the fact that in a market-oriented economy, firms will hire and fire workers. Nor is there much to be done about how the evolving age structure of the economy, or unexpected shifts in productivity, will affect the natural rate of unemployment for a time. However, as the example of high ongoing unemployment rates for many European countries illustrates, government policy clearly can affect the natural rate of unemployment that will persist even when GDP is growing.

When a government enacts policies that will affect workers or employers, it must examine how these policies will affect the information and incentives employees and employers have to seek each other out. For example, the government may have a role to play in helping some of the unemployed with job searches. The design of government programs that offer assistance to unemployed workers and protections to employed workers may need to be rethought so that they will not unduly discourage the supply of labor. Similarly, rules that make it difficult for businesses to begin or to expand may need to be redesigned so that they will not unduly discourage the demand for labor. The message is not that all laws affecting labor markets should be repealed, but only that when such laws are enacted, a society that cares about unemployment will need to consider the tradeoffs involved.

Note:**The Mysterious Case of the Missing Candidates**

After reading the chapter you might think the current unemployment conundrum may be due to structural unemployment. Indeed, there is a mismatch between the skills employers are seeking and the skills the unemployed possess. But Peter Cappelli has a slightly different view on this—it is called the purple squirrel. The what?

In human resource parlance, a purple squirrel is a job candidate who is a perfect fit for all of the many different responsibilities of a position. A purple squirrel candidate could step into a multi-faceted position with no training and permit the firm to higher fewer people because the worker is

so versatile. During the Great Recession, Human Resources (HR) positions were reduced. This means today's hiring managers are drafting job descriptions and requirements without much, if any HR feedback. "It turns out it's typically the case that employers' requirements are crazy, they're not paying enough, or their applicant screening is so rigid that nobody gets through," Cappelli stated in a 2012 Knowledge@Wharton interview about the findings in his book, *Why Good People Can't Find Jobs: Chasing After the Purple Squirrel*. In short, managers are searching for "purple squirrels" when what they really need are just versatile workers. There really is not a shortage of "normal squirrels"—candidates who are versatile workers. The managers just cannot find them because their requirements, screening processes, and compensation will filter out all but the "purple" ones.

Key Concepts and Summary

The natural rate of unemployment is the rate of unemployment that would be caused by the economic, social, and political forces in the economy even when the economy is not in a recession. These factors include the frictional unemployment that occurs when people are put out of work for a time by the shifts of a dynamic and changing economy and any laws concerning conditions of hiring and firing have the undesired side effect of discouraging job formation. They also include structural unemployment, which occurs when demand shifts permanently away from a certain type of job skill.

Self-Check Questions

Exercise:

Problem:

Is the increase in labor force participation rates among women better thought of as causing an increase in cyclical unemployment or an increase in the natural rate of unemployment? Why?

Solution:

The increase in labor supply was a social demographic trend—it was not caused by the economy falling into a recession. Therefore, the influx of women into the work force increased the natural rate of unemployment.

Exercise:

Problem:

Many college students graduate from college before they have found a job. When graduates begin to look for a job, they are counted as what category of unemployed?

Solution:

New entrants to the labor force, whether from college or otherwise, are counted as frictionally unemployed until they find a job.

Review Questions

Exercise:

Problem:

What term describes the remaining level of unemployment that occurs even when the economy is healthy?

Exercise:

Problem:

What forces create the natural rate of unemployment for an economy?

Exercise:

Problem:

Would you expect the natural rate of unemployment to be roughly the same in different countries?

Exercise:

Problem:

Would you expect the natural rate of unemployment to remain the same within one country over the long run of several decades?

Exercise:

Problem:

What is frictional unemployment? Give examples of frictional unemployment.

Exercise:

Problem:

What is structural unemployment? Give examples of structural unemployment.

Exercise:

Problem:

After several years of economic growth, would you expect the unemployment in an economy to be mainly cyclical or mainly due to the natural rate of unemployment? Why?

Exercise:

Problem:

What type of unemployment (cyclical, frictional, or structural) applies to each of the following:

- A. landscapers laid off in response to drop in new housing construction during a recession.
- B. coal miners laid off due to EPA regulations that shut down coal fired power
- C. a financial analyst who quits his/her job in Chicago and is pursuing similar work in Arizona
- D. printers laid off due to drop in demand for printed catalogues and flyers as firms go the internet to promote and advertise their

products.

E. factory workers in the U.S. laid off as the plants shut down and move to Mexico and Ireland.

Critical Thinking Questions

Exercise:

Problem:

Under what condition would a decrease in unemployment be bad for the economy?

Exercise:

Problem:

Under what condition would an increase in the unemployment rate be a positive sign?

Exercise:

Problem:

As the baby boom generation retires, the ratio of retirees to workers will increase noticeably. How will this affect the Social Security program? How will this affect the standard of living of the average American?

Exercise:

Problem:

Unemployment rates have been higher in many European countries in recent decades than in the United States. Is the main reason for this long-term difference in unemployment rates more likely to be cyclical unemployment or the natural rate of unemployment? Explain briefly.

Exercise:

Problem:

Is it desirable to pursue a goal of zero unemployment? Why or why not?

Exercise:**Problem:**

Is it desirable to eliminate natural unemployment? Why or why not?
Hint: Think about what our economy would look like today and what assumptions would have to be met to have a zero rate of natural unemployment.

Exercise:**Problem:**

The U.S. unemployment rate increased from 4.6% in July 2001 to 5.9% by June 2002. Without studying the subject in any detail, would you expect that a change of this kind is more likely to be due to cyclical unemployment or a change in the natural rate of unemployment? Why?

Problems**Exercise:****Problem:**

As the baby boomer generation retires, what should happen to wages and employment? Can you show this graphically?

References

Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey. *Accessed March 6, 2015*
<http://data.bls.gov/timeseries/LNS14000000>.

Cappelli, P. (20 June 2012). “Why Good People Can’t Get Jobs: Chasing After the Purple Squirrel.”

<http://knowledge.wharton.upenn.edu/article.cfm?articleid=3027>.

Glossary

frictional unemployment

unemployment that occurs as workers move between jobs

natural rate of unemployment

the unemployment rate that would exist in a growing and healthy economy from the combination of economic, social, and political factors that exist at a given time

structural unemployment

unemployment that occurs because individuals lack skills valued by employers

Introduction to Inflation
class="introduction"
Big Bucks in Zimbabwe

This bill was worth
100 billion
Zimbabwean dollars
when issued in 2008.
There were even bills
issued with a face
value of 100 trillion
Zimbabwean dollars.
The bills had
\$100,000,000,000,000
written on them.
Unfortunately, they
were almost
worthless. At one
point, 621,984,228
Zimbabwean dollars
were equal to one
U.S. dollar.
Eventually, the
country abandoned its
own currency and
allowed foreign
currency to be used
for purchases. (Credit:
modification of work
by Samantha
Marx/Flickr Creative
Commons)

**Note:****A \$550 Million Loaf of Bread?**

If you were born within the last three decades in the United States, Canada, or many other countries in the developed world, you probably have no real experience with a high rate of inflation. Inflation is when most prices in an entire economy are rising. But there is an extreme form of inflation called hyperinflation. This occurred in Germany between 1921 and 1928, and more recently in Zimbabwe between 2008 and 2009. In November of 2008, Zimbabwe had an inflation rate of 79.6 billion percent. In contrast, in 2014, the United States had an average annual rate of inflation of 1.6%. Zimbabwe's inflation rate was so high it is difficult to comprehend. So, let's put it into context. It is equivalent to price increases of 98% per day. This means that, from one day to the next, prices essentially double. What is life like in an economy afflicted with hyperinflation? Not like anything you are familiar with. Prices for commodities in Zimbabwean dollars were adjusted several times *each day*. There was no desire to hold on to currency since it lost value by the minute. The people there spent a great deal of time getting rid of any cash they acquired by purchasing whatever food or other commodities they could find. At one point, a loaf of bread cost 550 million Zimbabwean dollars. Teachers were paid in the trillions a month; however this was equivalent to only one U.S. dollar a day. At its

height, it took 621,984,228 Zimbabwean dollars to purchase one U.S. dollar.

Government agencies had no money to pay their workers so they started printing money to pay their bills rather than raising taxes. Rising prices caused the government to enact price controls on private businesses, which led to shortages and the emergence of black markets. In 2009, the country abandoned its currency and allowed foreign currencies to be used for purchases.

How does this happen? How can both government and the economy fail to function at the most basic level? Before we consider these extreme cases of hyperinflation, let's first look at inflation itself.

Note:

Introduction to Inflation

In this chapter, you will learn about:

- Tracking Inflation
- How Changes in the Cost of Living are Measured
- How the U.S. and Other Countries Experience Inflation
- The Confusion Over Inflation
- Indexing and Its Limitations

Inflation is a general and ongoing rise in the level of prices in an entire economy. Inflation does not refer to a change in relative prices. A relative price change occurs when you see that the price of tuition has risen, but the price of laptops has fallen. Inflation, on the other hand, means that there is pressure for prices to rise in most markets in the economy. In addition, price increases in the supply-and-demand model were one-time events, representing a shift from a previous equilibrium to a new one. Inflation implies an ongoing rise in prices. If inflation happened for one year and then stopped—well, then it would not be inflation any more.

This chapter begins by showing how to combine prices of individual goods and services to create a measure of overall inflation. It discusses the historical and recent experience of inflation, both in the United States and in other countries around the world. Other chapters have sometimes included a note under an exhibit or a parenthetical reminder in the text saying that the numbers have been adjusted for inflation. In this chapter, it is time to show how to use inflation statistics to adjust other economic variables, so that you can tell how much of, say, the rise in GDP over different periods of time can be attributed to an actual increase in the production of goods and services and how much should be attributed to the fact that prices for most things have risen.

Inflation has consequences for people and firms throughout the economy, in their roles as lenders and borrowers, wage-earners, taxpayers, and consumers. The chapter concludes with a discussion of some imperfections and biases in the inflation statistics, and a preview of policies for fighting inflation that will be discussed in other chapters.

Tracking Inflation

By the end of this section, you will be able to:

- Calculate the annual rate of inflation
- Explain and use index numbers and base years when simplifying the total quantity spent over a year for products
- Calculate inflation rates using index numbers

Dinner table conversations where you might have heard about inflation usually entail reminiscing about when “everything seemed to cost so much less. You used to be able to buy three gallons of gasoline for a dollar and then go see an afternoon movie for another dollar.” [\[link\]](#) compares some prices of common goods in 1970 and 2014. Of course, the average prices shown in this table may not reflect the prices where you live. The cost of living in New York City is much higher than in Houston, Texas, for example. In addition, certain products have evolved over recent decades. A new car in 2014, loaded with antipollution equipment, safety gear, computerized engine controls, and many other technological advances, is a more advanced machine (and more fuel efficient) than your typical 1970s car. However, put details like these to one side for the moment, and look at the overall pattern. The primary reason behind the price rises in [\[link\]](#)—and all the price increases for the other products in the economy—is not specific to the market for housing or cars or gasoline or movie tickets. Instead, it is part of a general rise in the level of all prices. In 2014, \$1 had about the same purchasing power in overall terms of goods and services as 18 cents did in 1972, because of the amount of inflation that has occurred over that time period.

Items	1970	2014
Pound of ground beef	\$0.66	\$4.16
Pound of butter	\$0.87	\$2.93
Movie ticket	\$1.55	\$8.17
Sales price of new home (median)	\$22,000	\$280,000
New car	\$3,000	\$32,531
Gallon of gasoline	\$0.36	\$3.36
Average hourly wage for a manufacturing worker	\$3.23	\$19.55
Per capita GDP	\$5,069	\$53,041.98

Price Comparisons, 1970 and 2014(Sources: See chapter References at end of book.)

Moreover, the power of inflation does not affect just goods and services, but wages and income levels, too. The second-to-last row of [\[link\]](#) shows that the average hourly wage for a manufacturing worker increased nearly six-fold from 1970 to 2014. Sure, the average worker in 2014 is better educated and more productive than the average worker in 1970—but not six times more productive. Sure, per capita GDP increased substantially from 1970 to 2014, but is the average person in the U.S. economy really more than eight times better off in just 44 years? Not likely.

A modern economy has millions of goods and services whose prices are continually quivering in the breezes of supply and demand. How can all of these shifts in price be boiled down to a single inflation rate? As with many problems in economic measurement, the conceptual answer is reasonably straightforward: Prices of a variety of

goods and services are combined into a single price level; the inflation rate is simply the percentage change in the price level. Applying the concept, however, involves some practical difficulties.

The Price of a Basket of Goods

To calculate the price level, economists begin with the concept of a **basket of goods and services**, consisting of the different items individuals, businesses, or organizations typically buy. The next step is to look at how the prices of those items change over time. In thinking about how to combine individual prices into an overall price level, many people find that their first impulse is to calculate the average of the prices. Such a calculation, however, could easily be misleading because some products matter more than others.

Changes in the prices of goods for which people spend a larger share of their incomes will matter more than changes in the prices of goods for which people spend a smaller share of their incomes. For example, an increase of 10% in the rental rate on housing matters more to most people than whether the price of carrots rises by 10%. To construct an overall measure of the price level, economists compute a weighted average of the prices of the items in the basket, where the weights are based on the actual quantities of goods and services people buy. The following Work It Out feature walks you through the steps of calculating the annual rate of inflation based on a few products.

Note:
Calculating an Annual Rate of Inflation
Consider the simple basket of goods with only three items, represented in [\[link\]](#). Say that in any given month, a college student spends money on 20 hamburgers, one bottle of aspirin, and five movies. Prices for these items over four years are given in the table through each time period (Pd). Prices of some goods in the basket may rise while others fall. In this example, the price of aspirin does not change over the four years, while movies increase in price and hamburgers bounce up and down. Each year, the cost of buying the given basket of goods at the prices prevailing at that time is shown.

Items	Hamburger	Aspirin	Movies	Total	Inflation Rate
Qty	20	1 bottle	5	-	-
(Pd 1) Price	\$3.00	\$10.00	\$6.00	-	-
(Pd 1) Amount Spent	\$60.00	\$10.00	\$30.00	\$100.00	-
(Pd 2) Price	\$3.20	\$10.00	\$6.50	-	-
(Pd 2) Amount Spent	\$64.00	\$10.00	\$32.50	\$106.50	6.5%
(Pd 3) Price	\$3.10	\$10.00	\$7.00	-	-
(Pd 3) Amount Spent	\$62.00	\$10.00	\$35.00	\$107.00	0.5%
(Pd 4) Price	\$3.50	\$10.00	\$7.50	-	-
(Pd 4) Amount Spent	\$70.00	\$10.00	\$37.50	\$117.50	9.8%

A College Student’s Basket of Goods

To calculate the annual rate of inflation in this example:

Step 1. Find the percentage change in the cost of purchasing the overall basket of goods between the time periods. The general equation for percentage changes between two years, whether in the context of inflation or in any other calculation, is:

Equation:

$$\frac{(\text{Level in new year} - \text{Level in previous year})}{\text{Level in previous year}} = \text{Percentage change}$$

Step 2. From period 1 to period 2, the total cost of purchasing the basket of goods in [\[link\]](#) rises from \$100 to \$106.50. Therefore, the percentage change over this time—the inflation rate—is:

Equation:

$$\frac{(106.50 - 100)}{100.0} = 0.065 = 6.5\%$$

Step 3. From period 2 to period 3, the overall change in the cost of purchasing the basket rises from \$106.50 to \$107. Thus, the inflation rate over this time, again calculated by the percentage change, is approximately:

Equation:

$$\frac{(107 - 106.50)}{106.50} = 0.0047 = 0.47\%$$

Step 4. From period 3 to period 4, the overall cost rises from \$107 to \$117.50. The inflation rate is thus:

Equation:

$$\frac{(117.50 - 107)}{107} = 0.098 = 9.8\%$$

This calculation of the change in the total cost of purchasing a basket of goods takes into account how much is spent on each good. Hamburgers are the lowest-priced good in this example, and aspirin is the highest-priced. If an individual buys a greater quantity of a low-price good, then it makes sense that changes in the price of that good should have a larger impact on the buying power of that person's money. The larger impact of hamburgers shows up in the "amount spent" row, where, in all time periods, hamburgers are the largest item within the amount spent row.

Index Numbers

The numerical results of a calculation based on a basket of goods can get a little messy. The simplified example in [\[link\]](#) has only three goods and the prices are in even dollars, not numbers like 79 cents or \$124.99. If the list of products was much longer, and more realistic prices were used, the total quantity spent over a year might be some messy-looking number like \$17,147.51 or \$27,654.92.

To simplify the task of interpreting the price levels for more realistic and complex baskets of goods, the price level in each period is typically reported as an **index number**, rather than as the dollar amount for buying the basket of goods. Price indices are created to calculate an overall average change in relative prices over time. To convert the money spent on the basket to an index number, economists arbitrarily choose one year to be the **base year**, or starting point from which we measure changes in prices. The base year, by definition, has an index number equal to 100. This sounds complicated, but it is really a simple math trick. In the example above, say that time period 3 is chosen as the base year. Since the total amount of spending in that year is \$107, we divide that amount by itself (\$107) and multiply by 100. Mathematically, that is equivalent to dividing \$107 by 100, or \$1.07. Doing either will give us an index in the base year of 100. Again, this is because the index number in the base year *always* has to have a value of 100. Then, to figure out the values of the index number for the other years, we divide the dollar amounts for the other years by 1.07 as well. Note also that the dollar signs cancel out so that index numbers have no units.

Calculations for the other values of the index number, based on the example presented in [\[link\]](#) are shown in [\[link\]](#). Because the index numbers are calculated so that they are in exactly the same proportion as the total dollar cost of

purchasing the basket of goods, the inflation rate can be calculated based on the index numbers, using the percentage change formula. So, the inflation rate from period 1 to period 2 would be

Equation:

$$\frac{(99.5 - 93.4)}{93.4} = 0.065 = 6.5\%$$

This is the same answer that was derived when measuring inflation based on the dollar cost of the basket of goods for the same time period.

	Total Spending	Index Number	Inflation Rate Since Previous Period
Period 1	\$100	$\frac{100}{1.07} = 93.4$	
Period 2	\$106.50	$\frac{106.50}{1.07} = 99.5$	$\frac{(99.5 - 93.4)}{93.4} = 0.065 = 6.5\%$
Period 3	\$107	$\frac{107}{1.07} = 100.0$	$\frac{100 - 99.5}{99.5} = 0.005 = 0.5\%$
Period 4	\$117.50	$\frac{117.50}{1.07} = 109.8$	$\frac{109.8 - 100}{100} = 0.098 = 9.8\%$

Calculating Index Numbers When Period 3 is the Base Year

If the inflation rate is the same whether it is based on dollar values or index numbers, then why bother with the index numbers? The advantage is that indexing allows easier eyeballing of the inflation numbers. If you glance at two index numbers like 107 and 110, you know automatically that the rate of inflation between the two years is about, but not quite exactly equal to, 3%. By contrast, imagine that the price levels were expressed in absolute dollars of a large basket of goods, so that when you looked at the data, the numbers were \$19,493.62 and \$20,009.32. Most people find it difficult to eyeball those kinds of numbers and say that it is a change of about 3%. However, the two numbers expressed in absolute dollars are exactly in the same proportion of 107 to 110 as the previous example. If you're wondering why simple subtraction of the index numbers wouldn't work, read the following Clear It Up feature.

Note:

Why do you not just subtract index numbers?

A word of warning: When a price index moves from, say, 107 to 110, the rate of inflation is not *exactly* 3%. Remember, the inflation rate is not derived by subtracting the index numbers, but rather through the percentage-change calculation. The precise inflation rate as the price index moves from 107 to 110 is calculated as $(110 - 107) / 107 = 0.028 = 2.8\%$. When the base year is fairly close to 100, a quick subtraction is not a terrible shortcut to calculating the inflation rate—but when precision matters down to tenths of a percent, subtracting will not give the right answer.

Two final points about index numbers are worth remembering. First, index numbers have no dollar signs or other units attached to them. Although index numbers can be used to calculate a percentage inflation rate, the index numbers themselves do not have percentage signs. Index numbers just mirror the proportions found in other data. They transform the other data so that the data are easier to work with.

Second, the choice of a base year for the index number—that is, the year that is automatically set equal to 100—is arbitrary. It is chosen as a starting point from which changes in prices are tracked. In the official inflation statistics, it is common to use one base year for a few years, and then to update it, so that the base year of 100 is relatively close to the present. But any base year that is chosen for the index numbers will result in exactly the same inflation rate. To see this in the previous example, imagine that period 1, when total spending was \$100, was also chosen as the base year, and given an index number of 100. At a glance, you can see that the index numbers would now exactly match the dollar figures, the inflation rate in the first period would be 6.5%, and so on.

Now that we see how indexes work to track inflation, the next module will show us how the cost of living is measured.

Note:

Watch this [video](#) from the cartoon *Duck Tales* to view a mini-lesson on inflation.



Key Concepts and Summary

The price level is measured by using a basket of goods and services and calculating how the total cost of buying that basket of goods will increase over time. The price level is often expressed in terms of index numbers, which transform the cost of buying the basket of goods and services into a series of numbers in the same proportion to each other, but with an arbitrary base year of 100. The rate of inflation is measured as the percentage change between price levels or index numbers over time.

Self-Check Questions

Exercise:

Problem:

[\[link\]](#) shows the prices of fruit purchased by the typical college student from 2001 to 2004. What is the amount spent each year on the “basket” of fruit with the quantities shown in column 2?

Items	Qty	(2001) Price	(2001) Amount Spent	(2002) Price	(2002) Amount Spent	(2003) Price	(2003) Amount Spent	(2004) Price
Apples	10	\$0.50		\$0.75		\$0.85		\$0.85
Bananas	12	\$0.20		\$0.25		\$0.25		\$0.25
Grapes	2	\$0.65		\$0.70		\$0.90		\$0.90

Items	Qty	(2001) Price	(2001) Amount Spent	(2002) Price	(2002) Amount Spent	(2003) Price	(2003) Amount Spent	(2004) Price
Raspberries	1	\$2.00		\$1.90		\$2.05		\$2.15
Total								

Solution:

To compute the amount spent on each fruit in each year, you multiply the quantity of each fruit by the price.

- 10 apples \times 50 cents each = \$5.00 spent on apples in 2001.
- 12 bananas \times 20 cents each = \$2.40 spent on bananas in 2001.
- 2 bunches of grapes at 65 cents each = \$1.30 spent on grapes in 2001.
- 1 pint of raspberries at \$2 each = \$2.00 spent on raspberries in 2001.

Adding up the amounts gives you the total cost of the fruit basket. The total cost of the fruit basket in 2001 was $\$5.00 + \$2.40 + \$1.30 + \$2.00 = \$10.70$. The total costs for all the years are shown in the following table.

2001	2002	2003	2004
\$10.70	\$13.80	\$15.35	\$16.31

Exercise:

Problem: Construct the price index for a “fruit basket” in each year using 2003 as the base year.

Solution:

If 2003 is the base year, then the index number has a value of 100 in 2003. To transform the cost of a fruit basket each year, we divide each year’s value by \$15.35, the value of the base year, and then multiply the result by 100. The price index is shown in the following table.

2001	2002	2003	2004
69.71	89.90	100.00	106.3

Note that the base year has a value of 100; years before the base year have values less than 100; and years after have values more than 100.

Exercise:

Problem: Compute the inflation rate for fruit prices from 2001 to 2004.

Solution:

The inflation rate is calculated as the percentage change in the price index from year to year. For example, the inflation rate between 2001 and 2002 is $(84.61 - 69.71) / 69.71 = 0.2137 = 21.37\%$. The inflation rates for all the years are shown in the last row of the following table, which includes the two previous answers.

Items	Qty	(2001) Price	(2001) Amount Spent	(2002) Price	(2002) Amount Spent	(2003) Price	(2003) Amount Spent	(2004) Price
Apples	10	\$0.50	\$5.00	\$0.75	\$7.50	\$0.85	\$8.50	\$0.85
Bananas	12	\$0.20	\$2.40	\$0.25	\$3.00	\$0.25	\$3.00	\$0.25
Grapes	2	\$0.65	\$1.30	\$0.70	\$1.40	\$0.90	\$1.80	\$0.95
Raspberries	1	\$2.00	\$2.00	\$1.90	\$1.90	\$2.05	\$2.05	\$2.15
Total			\$10.70		\$13.80		\$15.35	
Price Index			69.71		84.61		100.00	
Inflation Rate					21.37%		18.19%	

Exercise:

Problem:

Edna is living in a retirement home where most of her needs are taken care of, but she has some discretionary spending. Based on the basket of goods in [\[link\]](#), by what percentage does Edna's cost of living increase between time 1 and time 2?

Items	Quantity	(Time 1) Price	(Time 2) Price
Gifts for grandchildren	12	\$50	\$60
Pizza delivery	24	\$15	\$16
Blouses	6	\$60	\$50
Vacation trips	2	\$400	\$420

Solution:

Begin by calculating the total cost of buying the basket in each time period, as shown in the following table.

Items	Quantity	(Time 1) Price	(Time 1) Total Cost	(Time 2) Price	(Time 2) Total Cost
Gifts	12	\$50	\$600	\$60	\$720
Pizza	24	\$15	\$360	\$16	\$384
Blouses	6	\$60	\$360	\$50	\$300
Trips	2	\$400	\$800	\$420	\$840
Total Cost			\$2,120		\$2,244

The rise in cost of living is calculated as the percentage increase:

$$(2244 - 2120) / 2120 = 0.0585 = 5.85\%.$$

Review Questions**Exercise:**

Problem: How is a basket of goods and services used to measure the price level?

Exercise:

Problem: Why are index numbers used to measure the price level rather than dollar value of goods?

Exercise:

Problem: What is the difference between the price level and the rate of inflation?

Critical Thinking Question**Exercise:****Problem:**

Inflation rates, like most statistics, are imperfect measures. Can you identify some ways that the inflation rate for fruit does not perfectly capture the rising price of fruit?

Problems**Exercise:**

Problem:

The index number representing the price level changes from 110 to 115 in one year, and then from 115 to 120 the next year. Since the index number increases by five each year, is five the inflation rate each year? Is the inflation rate the same each year? Explain your answer.

Exercise:**Problem:**

The total price of purchasing a basket of goods in the United Kingdom over four years is: year 1=£940, year 2=£970, year 3=£1000, and year 4=£1070. Calculate two price indices, one using year 1 as the base year (set equal to 100) and the other using year 4 as the base year (set equal to 100). Then, calculate the inflation rate based on the first price index. If you had used the other price index, would you get a different inflation rate? If you are unsure, do the calculation and find out.

References

Sources for [\[link\]](#):

http://www.eia.gov/dnav/pet/pet_pri_gnd_a_epmr_pte_dpgal_w.htm

<http://data.bls.gov/cgi-bin/surveymost?ap>

<http://www.bls.gov/ro3/apmw.htm>

<http://www.autoblog.com/2014/03/12/who-can-afford-the-average-car-price-only-folks-in-washington/>

<https://www.census.gov/construction/nrs/pdf/uspricemon.pdf>

<http://www.bls.gov/news.release/empsit.t24.htm>

<http://variety.com/2015/film/news/movie-ticket-prices-increased-in-2014-1201409670/>

US Inflation Calculator. "Historical Inflation Rates: 1914-2013." Accessed March 4, 2015.

<http://www.usinflationcalculator.com/inflation/historical-inflation-rates/>.

Glossary**base year**

arbitrary year whose value as an index number is defined as 100; inflation from the base year to other years can easily be seen by comparing the index number in the other year to the index number in the base year—for example, 100; so, if the index number for a year is 105, then there has been exactly 5% inflation between that year and the base year

basket of goods and services

a hypothetical group of different items, with specified quantities of each one meant to represent a “typical” set of consumer purchases, used as a basis for calculating how the price level changes over time

index number

a unit-free number derived from the price level over a number of years, which makes computing inflation rates easier, since the index number has values around 100

inflation

a general and ongoing rise in the level of prices in an economy

How Changes in the Cost of Living are Measured

By the end of this section, you will be able to:

- Use the Consumer Price Index (CPI) to calculate U.S. inflation rates
- Identify several ways the Bureau of Labor Statistics avoids biases in the Consumer Price Index (CPI)
- Differentiate among the Consumer Price Index (CPI), the Producer Price Index (PPI), the International Price Index, the Employment Cost Index, and the GDP deflator.

The most commonly cited measure of inflation in the United States is the **Consumer Price Index (CPI)**. The CPI is calculated by government statisticians at the U.S. Bureau of Labor Statistics based on the prices in a fixed basket of goods and services that represents the purchases of the average family of four. In recent years, the statisticians have paid considerable attention to a subtle problem: that the change in the total cost of buying a fixed basket of goods and services over time is conceptually not quite the same as the change in the cost of living, because the cost of living represents how much it costs for a person to feel that his or her consumption provides an equal level of satisfaction or utility.

To understand the distinction, imagine that over the past 10 years, the cost of purchasing a fixed basket of goods increased by 25% and your salary also increased by 25%. Has your personal standard of living held constant? If you do not necessarily purchase an identical fixed basket of goods every year, then an inflation calculation based on the cost of a fixed basket of goods may be a misleading measure of how your cost of living has changed. Two problems arise here: substitution bias and quality/new goods bias.

When the price of a good rises, consumers tend to purchase less of it and to seek out substitutes instead. Conversely, as the price of a good falls, people will tend to purchase more of it. This pattern implies that goods with generally rising prices should tend over time to become less important in the overall basket of goods used to calculate inflation, while goods with falling prices should tend to become more important. Consider, as an example, a rise in the price of peaches by \$100 per pound. If consumers

were utterly inflexible in their demand for peaches, this would lead to a big rise in the price of food for consumers. Alternatively, imagine that people are utterly indifferent to whether they have peaches or other types of fruit. Now, if peach prices rise, people completely switch to other fruit choices and the average price of food does not change at all. A fixed and unchanging basket of goods assumes that consumers are locked into buying exactly the same goods, regardless of price changes—not a very likely assumption. Thus, **substitution bias**—the rise in the price of a fixed basket of goods over time—tends to overstate the rise in a consumer’s true cost of living, because it does not take into account that the person can substitute away from goods whose relative prices have risen.

The other major problem in using a fixed basket of goods as the basis for calculating inflation is how to deal with the arrival of improved versions of older goods or altogether new goods. Consider the problem that arises if a cereal is improved by adding 12 essential vitamins and minerals—and also if a box of the cereal costs 5% more. It would clearly be misleading to count the entire resulting higher price as inflation, because the new price is being charged for a product of higher (or at least different) quality. Ideally, one would like to know how much of the higher price is due to the quality change, and how much of it is just a higher price. The Bureau of Labor Statistics, which is responsible for the computation of the Consumer Price Index, must deal with these difficulties in adjusting for quality changes.

Note:

Visit this [website](#) to view a list of Ford car prices between 1909 and 1927. Consider how these prices compare to today’s models. Is the product today of a different quality?



A new product can be thought of as an extreme improvement in quality—from something that did not exist to something that does. However, the basket of goods that was fixed in the past obviously does not include new goods created since then. The basket of goods and services used in the Consumer Price Index (CPI) is revised and updated over time, and so new products are gradually included. But the process takes some time. For example, room air conditioners were widely sold in the early 1950s, but were not introduced into the basket of goods behind the Consumer Price Index until 1964. The VCR and personal computer were available in the late 1970s and widely sold by the early 1980s, but did not enter the CPI basket of goods until 1987. By 1996, there were more than 40 million cellular phone subscribers in the United States—but cell phones were not yet part of the CPI basket of goods. The parade of inventions has continued, with the CPI inevitably lagging a few years behind.

The arrival of new goods creates problems with respect to the accuracy of measuring inflation. The reason people buy new goods, presumably, is that the new goods offer better value for money than existing goods. Thus, if the price index leaves out new goods, it overlooks one of the ways in which the cost of living is improving. In addition, the price of a new good is often higher when it is first introduced and then declines over time. If the new good is not included in the CPI for some years, until its price is already lower, the CPI may miss counting this price decline altogether. Taking these arguments together, the **quality/new goods bias** means that the rise in the price of a fixed basket of goods over time tends to overstate the rise in a consumer's true cost of living, because it does not take into account how improvements in the quality of existing goods or the invention of new goods improves the standard of living. The following Clear It Up feature is a must-read on how the CPI is comprised and calculated.

Note:

How do U.S. government statisticians measure the Consumer Price Index? When the U.S. Bureau of Labor Statistics (BLS) calculates the Consumer Price Index, the first task is to decide on a basket of goods that is representative of the purchases of the average household. This is done by

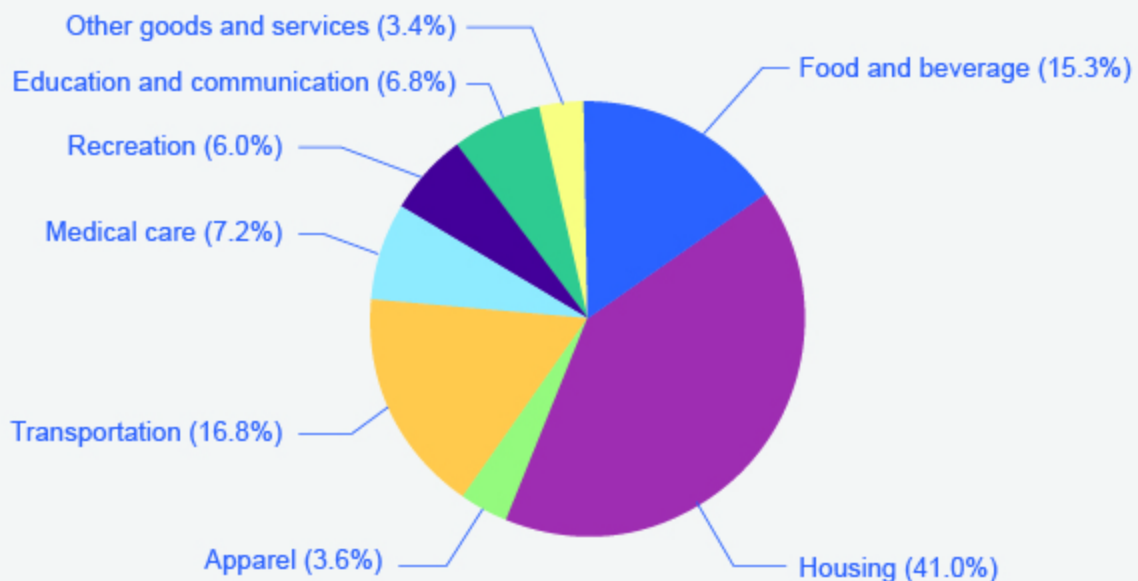
using the Consumer Expenditure Survey, a national survey of about 7,000 households, which provides detailed information on spending habits. Consumer expenditures are broken up into eight major groups, shown below, which in turn are broken up into more than 200 individual item categories. The BLS currently uses 1982–1984 as the base period. For each of the 200 individual expenditure items, the BLS chooses several hundred very specific examples of that item and looks at the prices of those examples. So, in figuring out the “breakfast cereal” item under the overall category of “foods and beverages,” the BLS picks several hundred examples of breakfast cereal. One example might be the price of a 24-oz. box of a particular brand of cereal sold at a particular store. The specific products and sizes and stores chosen are statistically selected to reflect what people buy and where they shop. The basket of goods in the Consumer Price Index thus consists of about 80,000 products; that is, several hundred specific products in over 200 broad-item categories. About one-quarter of these 80,000 specific products are rotated out of the sample each year, and replaced with a different set of products. The next step is to collect data on prices. Data collectors visit or call about 23,000 stores in 87 urban areas all over the United States every month to collect prices on these 80,000 specific products. A survey of 50,000 landlords or tenants is also carried out to collect information about rents. The Consumer Price Index is then calculated by taking the 80,000 prices of individual products and combining them, using weights (as shown in [\[link\]](#)) determined by the quantities of these products that people buy and allowing for factors like substitution between goods and quality improvements, into price indices for the 200 or so overall items. Then, the price indices for the 200 items are combined into an overall Consumer Price Index. According the Consumer Price Index website, there are eight categories used by data collectors:

The Eight Major Categories in the Consumer Price Index

1. Food and beverages (breakfast cereal, milk, coffee, chicken, wine, full-service meals, and snacks)
2. Housing (renter’s cost of housing, homeowner’s cost of housing, fuel oil, bedroom furniture)
3. Apparel (men’s shirts and sweaters, women’s dresses, jewelry)

4. Transportation (new vehicles, airline fares, gasoline, motor vehicle insurance)
5. Medical care (prescription drugs and medical supplies, physicians' services, eyeglasses and eye care, hospital services)
6. Recreation (televisions, cable television, pets and pet products, sports equipment, admissions)
7. Education and communication (college tuition, postage, telephone services, computer software and accessories)
8. Other goods and services (tobacco and smoking products, haircuts and other personal services, funeral expenses)

The Weighting of CPI Components



Of the eight categories used to generate the Consumer Price Index, housing is the highest at 41%. The next highest category, transportation at 16.8%, is less than half the size of housing. Other goods and services, and apparel, are the lowest at 3.4% and 3.6%, respectively. (Source: www.bls.gov/cpi)

The CPI and Core Inflation Index

Imagine if you were driving a company truck across the country- you probably would care about things like the prices of available roadside food and motel rooms as well as the truck's operating condition. However, the manager of the firm might have different priorities. He would care mostly about the truck's on-time performance and much less so about the food you were eating and the places you were staying. In other words, the company manager would be paying attention to the production of the firm, while ignoring transitory elements that impacted you, but did not affect the company's bottom line.

In a sense, a similar situation occurs with regard to measures of inflation. As we've learned, CPI measures prices as they affect everyday household spending. Well, a **core inflation index** is typically calculated by taking the CPI and excluding volatile economic variables. In this way, economists have a better sense of the underlying trends in prices that affect the cost of living.

Examples of excluded variables include energy and food prices, which can jump around from month to month because of the weather. According to an article by Kent Bernhard, during Hurricane Katrina in 2005, a key supply point for the nation's gasoline was nearly knocked out. Gas prices quickly shot up across the nation, in some places up to 40 cents a gallon in one day. This was not the cause of an economic policy but rather a short-lived event until the pumps were restored in the region. In this case, the CPI that month would register the change as a cost of living event to households, but the core inflation index would remain unchanged. As a result, the Federal Reserve's decisions on interest rates would not be influenced. Similarly, droughts can cause world-wide spikes in food prices that, if temporary, do not affect the nation's economic capability.

As former Chairman of the Federal Reserve Ben Bernanke noted in 1999 about the core inflation index, "It provide(s) a better guide to monetary policy than the other indices, since it measures the more persistent underlying inflation rather than transitory influences on the price level." Bernanke also noted that it helps communicate that every inflationary shock need not be responded to by the Federal Reserve since some price changes are transitory and not part of a structural change in the economy.

In sum, both the CPI and the core inflation index are important, but serve different audiences. The CPI helps households understand their overall cost of living from month to month, while the core inflation index is a preferred gauge from which to make important government policy changes.

Practical Solutions for the Substitution and the Quality/New Goods Biases

By the early 2000s, the Bureau of Labor Statistics was using alternative mathematical methods for calculating the Consumer Price Index, more complicated than just adding up the cost of a fixed basket of goods, to allow for some substitution between goods. It was also updating the basket of goods behind the CPI more frequently, so that new and improved goods could be included more rapidly. For certain products, the BLS was carrying out studies to try to measure the quality improvement. For example, with computers, an economic study can try to adjust for changes in speed, memory, screen size, and other characteristics of the product, and then calculate the change in price after these product changes are taken into account. But these adjustments are inevitably imperfect, and exactly how to make these adjustments is often a source of controversy among professional economists.

By the early 2000s, the substitution bias and quality/new goods bias had been somewhat reduced, so that since then the rise in the CPI probably overstates the true rise in inflation by only about 0.5% per year. Over one or a few years, this is not much; over a period of a decade or two, even half of a percent per year compounds to a more significant amount. In addition, the CPI tracks prices from physical locations, and not at online sites like Amazon, where prices can be lower.

When measuring inflation (and other economic statistics, too), a tradeoff arises between simplicity and interpretation. If the inflation rate is calculated with a basket of goods that is fixed and unchanging, then the calculation of an inflation rate is straightforward, but the problems of substitution bias and quality/new goods bias will arise. However, when the basket of goods is allowed to shift and evolve to reflect substitution toward

lower relative prices, quality improvements, and new goods, the technical details of calculating the inflation rate grow more complex.

Additional Price Indices: PPI, GDP Deflator, and More

The basket of goods behind the Consumer Price Index represents an average hypothetical U.S. household, which is to say that it does not exactly capture anyone's personal experience. When the task is to calculate an average level of inflation, this approach works fine. What if, however, you are concerned about inflation experienced by a certain group, like the elderly, or the poor, or single-parent families with children, or Hispanic-Americans? In specific situations, a price index based on the buying power of the average consumer may not feel quite right.

This problem has a straightforward solution. If the Consumer Price Index does not serve the desired purpose, then invent another index, based on a basket of goods appropriate for the group of interest. Indeed, the Bureau of Labor Statistics publishes a number of experimental price indices: some for particular groups like the elderly or the poor, some for different geographic areas, and some for certain broad categories of goods like food or housing.

The BLS also calculates several price indices that are not based on baskets of consumer goods. For example, the **Producer Price Index (PPI)** is based on prices paid for supplies and inputs by producers of goods and services. It can be broken down into price indices for different industries, commodities, and stages of processing (like finished goods, intermediate goods, crude materials for further processing, and so on). There is an **International Price Index** based on the prices of merchandise that is exported or imported. An **Employment Cost Index** measures wage inflation in the labor market. The **GDP deflator**, measured by the Bureau of Economic Analysis, is a price index that includes all the components of GDP (that is, consumption plus investment plus government plus exports minus imports). Unlike the CPI, its baskets are not fixed but re-calculate what that year's GDP would have been worth using the base-year's prices. MIT's Billion Prices Project is a more recent alternative attempt to measure prices: data are collected online from retailers and then composed into an index that is compared to the CPI (Source: <http://bpp.mit.edu/usa/>).

What's the best measure of inflation? If concerned with the most accurate measure of inflation, use the GDP deflator as it picks up the prices of goods and services produced. However, it is not a good measure of cost of living as it includes prices of many products not purchased by households (for example, aircraft, fire engines, factory buildings, office complexes, and bulldozers). If one wants the most accurate measure of inflation as it impacts households, use the CPI, as it only picks up prices of products purchased by households. That is why the CPI is sometimes referred to as the cost-of-living index. As the Bureau of Labor Statistics states on its website: "The 'best' measure of inflation for a given application depends on the intended use of the data."

Key Concepts and Summary

Measuring price levels with a fixed basket of goods will always have two problems: the substitution bias, by which a fixed basket of goods does not allow for buying more of what is relatively less expensive and less of what is relatively more expensive; and the quality/new goods bias, by which a fixed basket cannot take into account improvements in quality and the advent of new goods. These problems can be reduced in degree—for example, by allowing the basket of goods to evolve over time—but they cannot be totally eliminated. The most commonly cited measure of inflation is the Consumer Price Index (CPI), which is based on a basket of goods representing what the typical consumer buys. The Core Inflation Index further breaks down the CPI by excluding volatile economic variables. Several price indices are not based on baskets of consumer goods. The GDP deflator is based on all the components of GDP. The Producer Price Index is based on prices of supplies and inputs bought by producers of goods and services. An Employment Cost Index measures wage inflation in the labor market. An International Price Index is based on the prices of merchandise that is exported or imported.

Self-Check Questions

Exercise:

Problem:

[How Changes in the Cost of Living are Measured](#) introduced a number of different price indices. Which price index would be best to use to adjust your paycheck for inflation?

Solution:

Since the CPI measures the prices of the goods and services purchased by the typical urban consumer, it measures the prices of things that people buy with their paycheck. For that reason, the CPI would be the best price index to use for this purpose.

Exercise:**Problem:**

The Consumer Price Index is subject to the substitution bias and the quality/new goods bias. Are the Producer Price Index and the GDP Deflator also subject to these biases? Why or why not?

Solution:

The PPI is subject to those biases for essentially the same reasons as the CPI is. The GDP deflator picks up prices of what is actually purchased that year, so there are no biases. That is the advantage of using the GDP deflator over the CPI.

Review Questions**Exercise:****Problem:**

Why does “substitution bias” arise if the inflation rate is calculated based on a fixed basket of goods?

Exercise:

Problem:

Why does the “quality/new goods bias” arise if the inflation rate is calculated based on a fixed basket of goods?

Critical Thinking Question**Exercise:****Problem:**

Given the federal budget deficit in recent years, some economists have argued that by adjusting Social Security payments for inflation using the CPI, Social Security is overpaying recipients. What is the argument being made, and do you agree or disagree with it?

Exercise:**Problem:**

Why is the GDP deflator not an accurate measure of inflation as it impacts a household?

Exercise:**Problem:**

Imagine that the government statisticians who calculate the inflation rate have been updating the basic basket of goods once every 10 years, but now they decide to update it every five years. How will this change affect the amount of substitution bias and quality/new goods bias?

Exercise:**Problem:**

Describe a situation, either a government policy situation, an economic problem, or a private sector situation, where using the CPI to convert from nominal to real would be more appropriate than using the GDP deflator.

Exercise:

Problem:

Describe a situation, either a government policy situation, an economic problem, or a private sector situation, where using the GDP deflator to convert from nominal to real would be more appropriate than using the CPI.

References

Bernhard, Kent. "Pump Prices Jump Across U.S. after Katrina." *NBC News*, September 1, 2005. http://www.nbcnews.com/id/9146363/ns/business-local_business/t/pump-prices-jump-across-us-after-katrina/#.U00kRfk7um4.

Wynne, Mark A. "Core Inflation, A Review of Some Conceptual Issues." *Federal Reserve Bank of St. Louis*. p. 209. Accessed April 14, 2014. <http://research.stlouisfed.org/publications/review/08/05/part2/Wynne.pdf>

Glossary

Consumer Price Index (CPI)

a measure of inflation calculated by U.S. government statisticians based on the price level from a fixed basket of goods and services that represents the purchases of the average consumer

core inflation index

a measure of inflation typically calculated by taking the CPI and excluding volatile economic variables such as food and energy prices to better measure the underlying and persistent trend in long-term prices

Employment Cost Index

a measure of inflation based on wages paid in the labor market

GDP deflator

a measure of inflation based on the prices of all the components of GDP

International Price Index

a measure of inflation based on the prices of merchandise that is exported or imported

Producer Price Index (PPI)

a measure of inflation based on prices paid for supplies and inputs by producers of goods and services

quality/new goods bias

inflation calculated using a fixed basket of goods over time tends to overstate the true rise in cost of living, because it does not take into account improvements in the quality of existing goods or the invention of new goods

substitution bias

an inflation rate calculated using a fixed basket of goods over time tends to overstate the true rise in the cost of living, because it does not take into account that the person can substitute away from goods whose prices rise by a lot

How the U.S. and Other Countries Experience Inflation

By the end of this section, you will be able to:

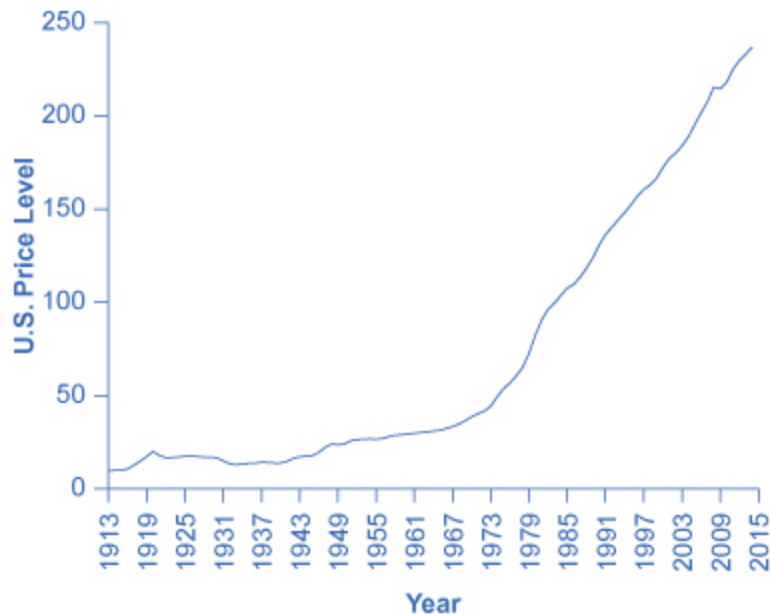
- Identify patterns of inflation for the United States using data from the Consumer Price Index
- Identify patterns of inflation on an international level

In the last three decades, inflation has been relatively low in the U.S. economy, with the Consumer Price Index typically rising 2% to 4% per year. Looking back over the twentieth century, there have been several periods where inflation caused the price level to rise at double-digit rates, but nothing has come close to hyperinflation.

Historical Inflation in the U.S. Economy

[\[link\]](#) (a) shows the level of prices in the Consumer Price Index stretching back to 1916. In this case, the base years (when the CPI is defined as 100) are set for the average level of prices that existed from 1982 to 1984. [\[link\]](#) (b) shows the annual percentage changes in the CPI over time, which is the inflation rate.

U.S. Price Level and Inflation Rates since 1913



(a) U.S. price level 1913-2014



(b) U.S. inflation rate 1913-2014

Graph a shows the trends in the U.S. price level from the year 1916 to 2014. In 1916, the graph starts out close to \$10, rises to around \$20 in 1920, stays around \$16 or \$17 until 1931, when it jumps to around \$15. It gradually increases, with periodic dips, until 2014, when it is around \$236. Graph b shows the trends in U.S. inflation rates from the year 1916 to 2014. In 1916, the graph starts out at 7.7%, jumps to close to

18% in 1917, drops drastically to close to –11% in 1921, goes up and down periodically, until settling to around 1.5% in 2014.

The first two waves of inflation are easy to characterize in historical terms: they are right after World War I and World War II. However, there are also two periods of severe negative inflation—called **deflation**—in the early decades of the twentieth century: one following the deep recession of 1920–21 and the other during the Great Depression of the 1930s. (Since inflation is a time when the buying power of money in terms of goods and services is reduced, deflation will be a time when the buying power of money in terms of goods and services increases.) For the period from 1900 to about 1960, the major inflations and deflations nearly balanced each other out, so the average annual rate of inflation over these years was only about 1% per year. A third wave of more severe inflation arrived in the 1970s and departed in the early 1980s.

Note:

Visit this [website](#) to use an inflation calculator and discover how prices have changed in the last 100 years.



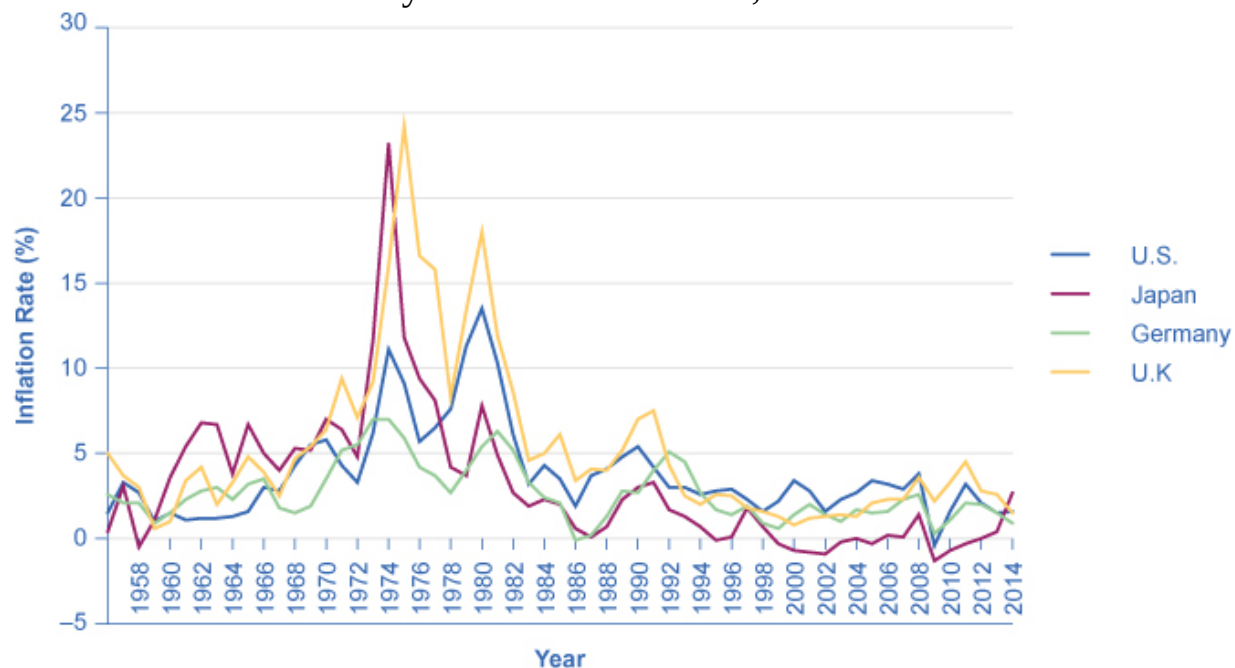
Times of recession or depression often seem to be times when the inflation rate is lower, as in the recession of 1920–1921, the Great Depression, the recession of 1980–1982, and the Great Recession in 2008–2009. There

were a few months in 2009 that were deflationary, but not at an annual rate. Recessions are typically accompanied by higher levels of unemployment, and the total demand for goods falls, pulling the price level down. Conversely, the rate of inflation often, but not always, seems to start moving up when the economy is growing very strongly, like right after wartime or during the 1960s. The frameworks for macroeconomic analysis, developed in other chapters, will explain why recession often accompanies higher unemployment and lower inflation, while rapid economic growth often brings lower unemployment but higher inflation.

Inflation around the World

Around the rest of the world, the pattern of inflation has been very mixed, as can be seen in [\[link\]](#) which shows inflation rates over the last several decades. Many industrialized countries, not just the United States, had relatively high inflation rates in the 1970s. For example, in 1975, Japan's inflation rate was over 8% and the inflation rate for the United Kingdom was almost 25%. In the 1980s, inflation rates came down in the United States and in Europe and have largely stayed down.

Countries with Relatively Low Inflation Rates, 1960–2014

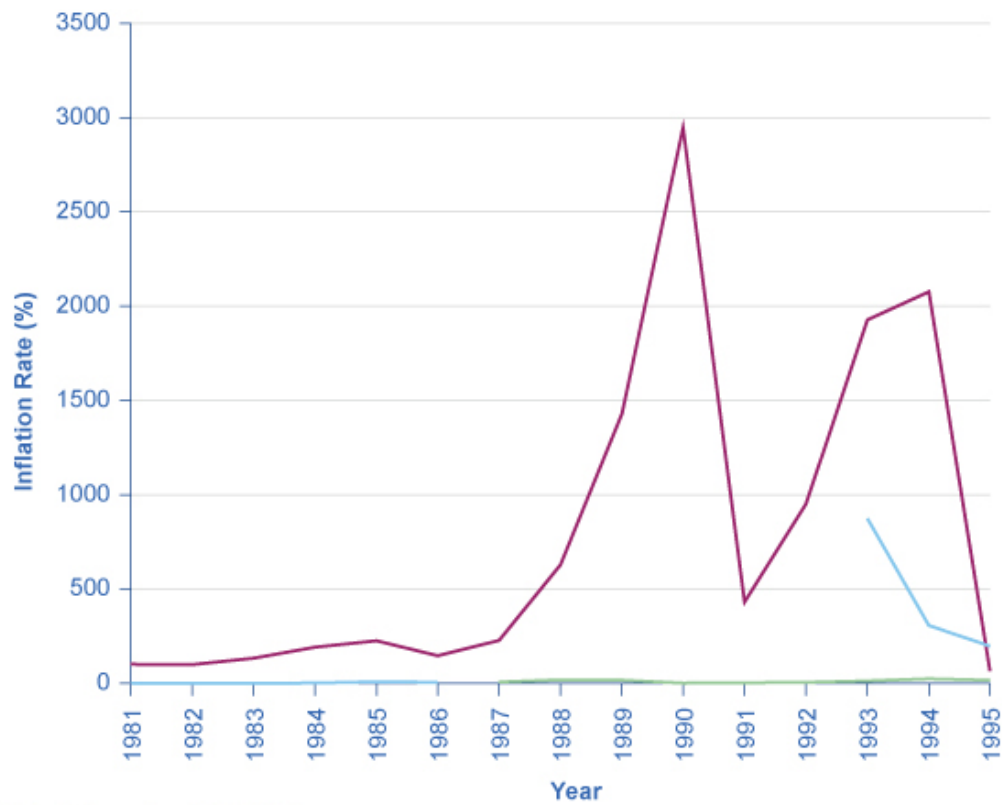


This chart shows the annual percentage change in consumer prices

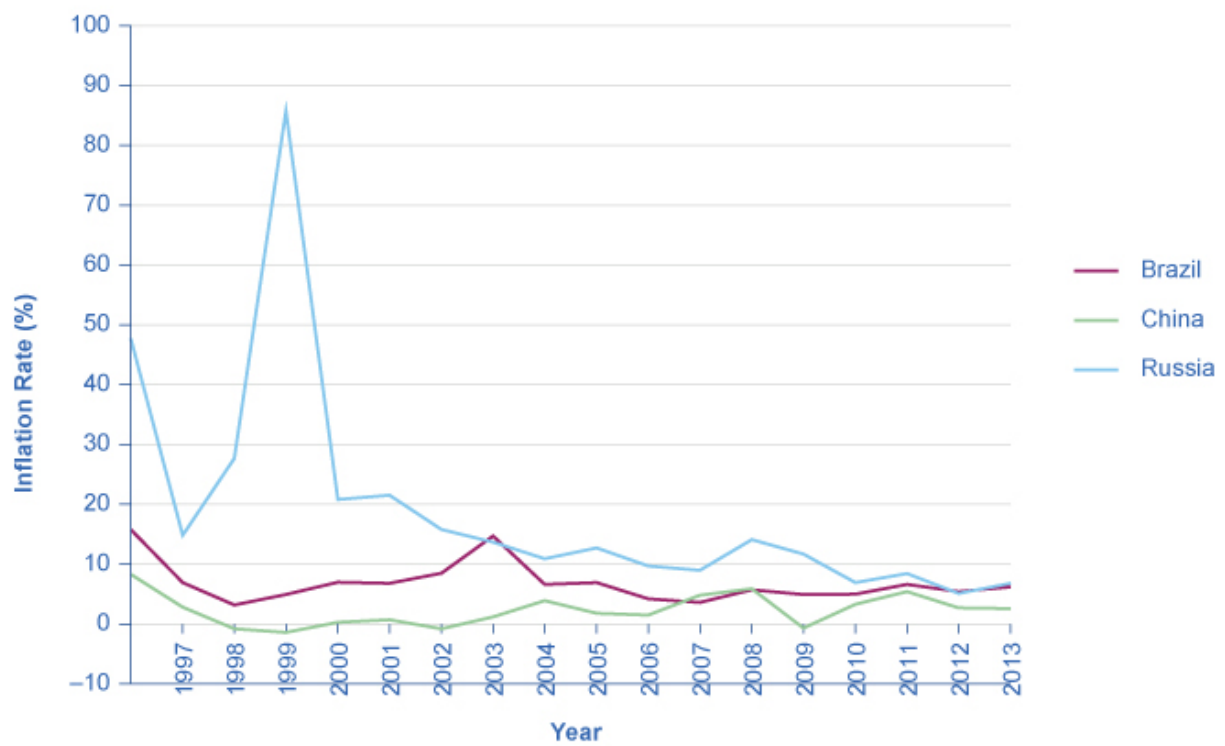
compared with the previous year's consumer prices in the United States, the United Kingdom, Japan, and Germany.

Countries with controlled economies in the 1970s, like the Soviet Union and China, historically had very low rates of measured inflation—because prices were forbidden to rise by law, except for the cases where the government deemed a price increase to be due to quality improvements. However, these countries also had perpetual shortages of goods, since forbidding prices to rise acts like a price ceiling and creates a situation where quantity demanded often exceeds quantity supplied. As Russia and China made a transition toward more market-oriented economies, they also experienced outbursts of inflation, although the statistics for these economies should be regarded as somewhat shakier. Inflation in China averaged about 10% per year for much of the 1980s and early 1990s, although it has dropped off since then. Russia experienced **hyperinflation**—an outburst of high inflation—of 2,500% per year in the early 1990s, although by 2006 Russia's consumer price inflation had dipped below 10% per year, as shown in [\[link\]](#). The closest the United States has ever gotten to hyperinflation was during the Civil War, 1860–1865, in the Confederate states.

Countries with Relatively High Inflation Rates, 1980–2013



(a) Inflation rates 1980-1995



(b) Inflation rates 1996-2013

These charts show the percentage change in consumer prices compared with the previous year's consumer prices in Brazil, China, and Russia. (a) Of these, Brazil and Russia experienced hyperinflation at some point between the mid-1980s and mid-1990s. (b) Though not as high, China and Nigeria also had high inflation rates in the mid-1990s. Even though their inflation rates have come down over the last two decades, several of these countries continue to see significant inflation rates. (Sources:

<http://research.stlouisfed.org/fred2/series/FPCPITOTLZGBRA>;
<http://research.stlouisfed.org/fred2/series/CHNCPIALLMINMEI>;
<http://research.stlouisfed.org/fred2/series/FPCPITOTLZGRUS>)

Many countries in Latin America experienced raging hyperinflation during the 1980s and early 1990s, with inflation rates often well above 100% per year. In 1990, for example, both Brazil and Argentina saw inflation climb above 2000%. Certain countries in Africa experienced extremely high rates of inflation, sometimes bordering on hyperinflation, in the 1990s. Nigeria, the most populous country in Africa, had an inflation rate of 75% in 1995.

In the early 2000s, the problem of inflation appears to have diminished for most countries, at least in comparison to the worst times of recent decades. As we noted in this earlier Bring it Home feature, in recent years, the world's worst example of hyperinflation was in Zimbabwe, where at one point the government was issuing bills with a face value of \$100 trillion (in Zimbabwean dollars)—that is, the bills had \$100,000,000,000,000 written on the front, but were almost worthless. In many countries, the memory of double-digit, triple-digit, and even quadruple-digit inflation is not very far in the past.

Key Concepts and Summary

In the U.S. economy, the annual inflation rate in the last two decades has typically been around 2% to 4%. The periods of highest inflation in the United States in the twentieth century occurred during the years after World

Wars I and II, and in the 1970s. The period of lowest inflation—actually, with deflation—was the Great Depression of the 1930s.

Self-Check Question

Exercise:

Problem:

Go to this [website](#) for the Purchasing Power Calculator at MeasuringWorth.com. How much money would it take today to purchase what one dollar would have bought in the year of your birth?

Solution:

The calculator requires you to input three numbers:

- The first year, in this case the year of your birth
- The amount of money you would want to translate in terms of its purchasing power
- The last year—now or the most recent year the calculator will accept

My birth year is 1955. The amount is \$1. The year 2012 is currently the latest year the calculator will accept. The simple purchasing power calculator shows that \$1 of purchases in 1955 would cost \$8.57 in 2012. The website also explains how the true answer is more complicated than that shown by the simple purchasing power calculator.

Review Questions

Exercise:

Problem:

What has been a typical range of inflation in the U.S. economy in the last decade or so?

Exercise:**Problem:**

Over the last century, during what periods was the U.S. inflation rate highest and lowest?

Exercise:

Problem: What is deflation?

Critical Thinking Question**Exercise:****Problem:**

Why do you think the U.S. experience with inflation over the last 50 years has been so much milder than in many other countries?

Problems**Exercise:****Problem:**

Within 1 or 2 percentage points, what has the U.S. inflation rate been during the last 20 years? Draw a graph to show the data.

Glossary

deflation

negative inflation; most prices in the economy are falling

hyperinflation

an outburst of high inflation that is often seen (although not exclusively) when economies shift from a controlled economy to a

market-oriented economy

The Confusion Over Inflation

By the end of this section, you will be able to:

- Explain how inflation can cause redistributions of purchasing power
- Identify ways inflation can blur the perception of supply and demand
- Explain the economic benefits and challenges of inflation

Economists usually oppose high inflation, but they oppose it in a milder way than many non-economists. Robert Shiller, one of 2013's Nobel Prize winners in economics, carried out several surveys during the 1990s about attitudes toward inflation. One of his questions asked, "Do you agree that preventing high inflation is an important national priority, as important as preventing drug abuse or preventing deterioration in the quality of our schools?" Answers were on a scale of 1–5, where 1 meant "Fully agree" and 5 meant "Completely disagree." For the U.S. population as a whole, 52% answered "Fully agree" that preventing high inflation was a highly important national priority and just 4% said "Completely disagree." However, among professional economists, only 18% answered "Fully agree," while the same percentage of 18% answered "Completely disagree."

The Land of Funny Money

What are the economic problems caused by inflation, and why do economists often regard them with less concern than the general public? Consider a very short story: "The Land of Funny Money."

One morning, everyone in the Land of Funny Money awakened to find that everything denominated in money had increased by 20%. The change was completely unexpected. Every price in every store was 20% higher. Paychecks were 20% higher. Interest rates were 20 % higher. The amount of money, everywhere from wallets to savings accounts, was 20% larger. This overnight inflation of prices made newspaper headlines everywhere in the Land of Funny Money. But the headlines quickly disappeared, as people realized that in terms of what they could actually buy with their incomes, this inflation had no economic impact. Everyone's pay could still buy exactly the same set of goods as it did before. Everyone's savings were still

sufficient to buy exactly the same car, vacation, or retirement that they could have bought before. Equal levels of inflation in all wages and prices ended up not mattering much at all.

When the people in Robert Shiller's surveys explained their concern about inflation, one typical reason was that they feared that as prices rose, they would not be able to afford to buy as much. In other words, people were worried because they did not live in a place like the Land of Funny Money, where all prices and wages rose simultaneously. Instead, people live here on Planet Earth, where prices might rise while wages do not rise at all, or where wages rise more slowly than prices.

Economists note that over most periods, the inflation level in prices is roughly similar to the inflation level in wages, and so they reason that, on average, over time, people's economic status is not greatly changed by inflation. If all prices, wages, and interest rates adjusted automatically and immediately with inflation, as in the Land of Funny Money, then no one's purchasing power, profits, or real loan payments would change. However, if other economic variables do not move exactly in sync with inflation, or if they adjust for inflation only after a time lag, then inflation can cause three types of problems: unintended redistributions of purchasing power, blurred price signals, and difficulties in long-term planning.

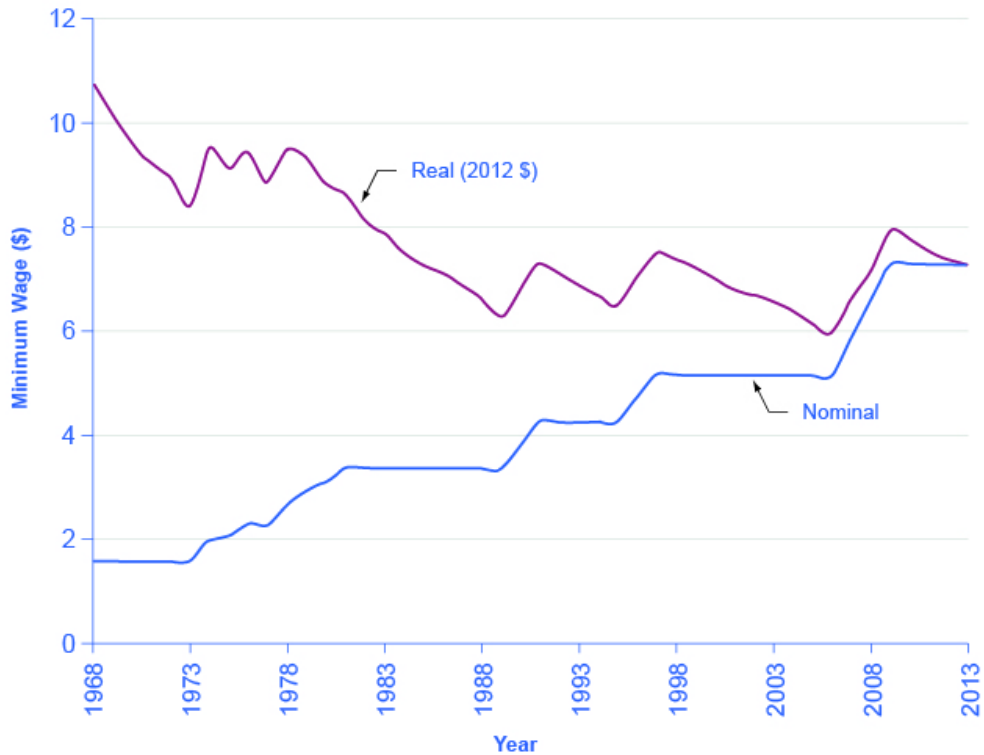
Unintended Redistributions of Purchasing Power

Inflation can cause redistributions of purchasing power that hurt some and help others. People who are hurt by inflation include those who are holding a lot of cash, whether it is in a safe deposit box or in a cardboard box under the bed. When inflation happens, the buying power of cash is diminished. But cash is only an example of a more general problem: anyone who has financial assets invested in a way that the nominal return does not keep up with inflation will tend to suffer from inflation. For example, if a person has money in a bank account that pays 4% interest, but inflation rises to 5%, then the real rate of return for the money invested in that bank account is negative 1%.

The problem of a good-looking nominal interest rate being transformed into an ugly-looking real interest rate can be worsened by taxes. The U.S. income tax is charged on the nominal interest received in dollar terms, without an adjustment for inflation. So, a person who invests \$10,000 and receives a 5% nominal rate of interest is taxed on the \$500 received—no matter whether the inflation rate is 0%, 5%, or 10%. If inflation is 0%, then the real interest rate is 5% and all \$500 is a gain in buying power. But if inflation is 5%, then the real interest rate is zero and the person had no real gain—but owes income tax on the nominal gain anyway. If inflation is 10%, then the real interest rate is *negative* 5% and the person is actually falling behind in buying power, but would still owe taxes on the \$500 in nominal gains.

Inflation can cause unintended redistributions for wage earners, too. Wages do typically creep up with inflation over time eventually. The last row of [\[link\]](#) at the start of this chapter showed that average hourly wage in the U.S. economy increased from \$3.23 in 1970 to \$19.55 in 2014, which is an increase by a factor of almost six. Over that time period, the Consumer Price Index increased by an almost identical amount. However, increases in wages may lag behind inflation for a year or two, since wage adjustments are often somewhat sticky and occur only once or twice a year. Moreover, the extent to which wages keep up with inflation creates insecurity for workers and may involve painful, prolonged conflicts between employers and employees. If the minimum wage is adjusted for inflation only infrequently, minimum wage workers are losing purchasing power from their nominal wages, as shown in [\[link\]](#).

U.S. Minimum Wage and Inflation



After adjusting for inflation, the federal minimum wage dropped more than 30 percent from 1967 to 2010, even though the nominal figure climbed from \$1.40 to \$7.25 per hour. Increases in the minimum wage in between 2008 and 2010 kept the decline from being worse—as it would have been if the wage had remained the same as it did from 1997 through 2007. (Sources: <http://www.dol.gov/whd/minwage/chart.htm>; <http://data.bls.gov/cgi-bin/surveymost?cu>)

One sizable group of people has often received a large share of their income in a form that does not increase over time: retirees who receive a private company pension. Most pensions have traditionally been set as a fixed nominal dollar amount per year at retirement. For this reason, pensions are called “defined benefits” plans. Even if inflation is low, the combination of inflation and a fixed income can create a substantial problem over time. A person who retires on a fixed income at age 65 will find that losing just 1%

to 2% of buying power per year to inflation compounds to a considerable loss of buying power after a decade or two.

Fortunately, pensions and other defined benefits retirement plans are increasingly rare, replaced instead by “defined contribution” plans, such as 401(k)s and 403(b)s. In these plans, the employer contributes a fixed amount to the worker’s retirement account on a regular basis (usually every pay check). The employee often contributes as well. The worker invests these funds in a wide range of investment vehicles. These plans are tax deferred, and they are portable so that if the individual takes a job with a different employer, their 401(k) comes with them. To the extent that the investments made generate real rates of return, retirees do not suffer from the inflation costs of traditional pensioners.

However, ordinary people can sometimes benefit from the unintended redistributions of inflation. Consider someone who borrows \$10,000 to buy a car at a fixed interest rate of 9%. If inflation is 3% at the time the loan is made, then the loan must be repaid at a real interest rate of 6%. But if inflation rises to 9%, then the real interest rate on the loan is zero. In this case, the borrower’s benefit from inflation is the lender’s loss. A borrower paying a fixed interest rate, who benefits from inflation, is just the flip side of an investor receiving a fixed interest rate, who suffers from inflation. The lesson is that when interest rates are fixed, rises in the rate of inflation tend to penalize suppliers of financial capital, who end up being repaid in dollars that are worth less because of inflation, while demanders of financial capital end up better off, because they can repay their loans in dollars that are worth less than originally expected.

The unintended redistributions of buying power caused by inflation may have a broader effect on society. America’s widespread acceptance of market forces rests on a perception that people’s actions have a reasonable connection to market outcomes. When inflation causes a retiree who built up a pension or invested at a fixed interest rate to suffer, however, while someone who borrowed at a fixed interest rate benefits from inflation, it is hard to believe that this outcome was deserved in any way. Similarly, when homeowners benefit from inflation because the price of their homes rises, while renters suffer because they are paying higher rent, it is hard to see any

useful incentive effects. One of the reasons that inflation is so disliked by the general public is a sense that it makes economic rewards and penalties more arbitrary—and therefore likely to be perceived as unfair – even dangerous, as the next Clear It Up feature shows.

Note:

Is there a connection between German hyperinflation and Hitler's rise to power?

Germany suffered an intense hyperinflation of its currency, the Mark, in the years after World War I, when the Weimar Republic in Germany resorted to printing money to pay its bills and the onset of the Great Depression created the social turmoil that Adolf Hitler could take advantage of in his rise to power. Shiller described the connection this way in a National Bureau of Economic Research 1996 Working Paper:

"A fact that is probably little known to young people today, even in Germany, is that the final collapse of the Mark in 1923, the time when the Mark's inflation reached astronomical levels (inflation of 35,974.9% in November 1923 alone, for an annual rate that month of $4.69 \times 10^{28}\%$), came in the same month as did Hitler's Beer Hall Putsch, his Nazi Party's armed attempt to overthrow the German government. This failed putsch resulted in Hitler's imprisonment, at which time he wrote his book *Mein Kampf*, setting forth an inspirational plan for Germany's future, suggesting plans for world domination. . . " " . . . Most people in Germany today probably do not clearly remember these events; this lack of attention to it may be because its memory is blurred by the more dramatic events that succeeded it (the Nazi seizure of power and World War II). However, to someone living through these historical events in sequence . . . [the putsch] may have been remembered as vivid evidence of the potential effects of inflation."

Blurred Price Signals

Prices are the messengers in a market economy, conveying information about conditions of demand and supply. Inflation blurs those price

messages. Inflation means that price signals are perceived more vaguely, like a radio program received with a lot of static. If the static becomes severe, it is hard to tell what is happening.

In Israel, when inflation accelerated to an annual rate of 500% in 1985, some stores stopped posting prices directly on items, since they would have had to put new labels on the items or shelves every few days to reflect inflation. Instead, a shopper just took items from a shelf and went up to the checkout register to find out the price for that day. Obviously, this situation makes comparing prices and shopping for the best deal rather difficult. When the levels and changes of prices become uncertain, businesses and individuals find it harder to react to economic signals. In a world where inflation is at a high rate, but bouncing up and down to some extent, does a higher price of a good mean that inflation has risen, or that supply of that good has decreased, or that demand for that good has increased? Should a buyer of the good take the higher prices as an economic hint to start substituting other products—or have the prices of the substitutes risen by an equal amount? Should a seller of the good take a higher price as a reason to increase production—or is the higher price only a sign of a general inflation in which the prices of all inputs to production are rising as well? The true story will presumably become clear over time, but at a given moment, who can say?

High and variable inflation means that the incentives in the economy to adjust in response to changes in prices are weaker. Markets will adjust toward their equilibrium prices and quantities more erratically and slowly, and many individual markets will experience a greater chance of surpluses and shortages.

Problems of Long-Term Planning

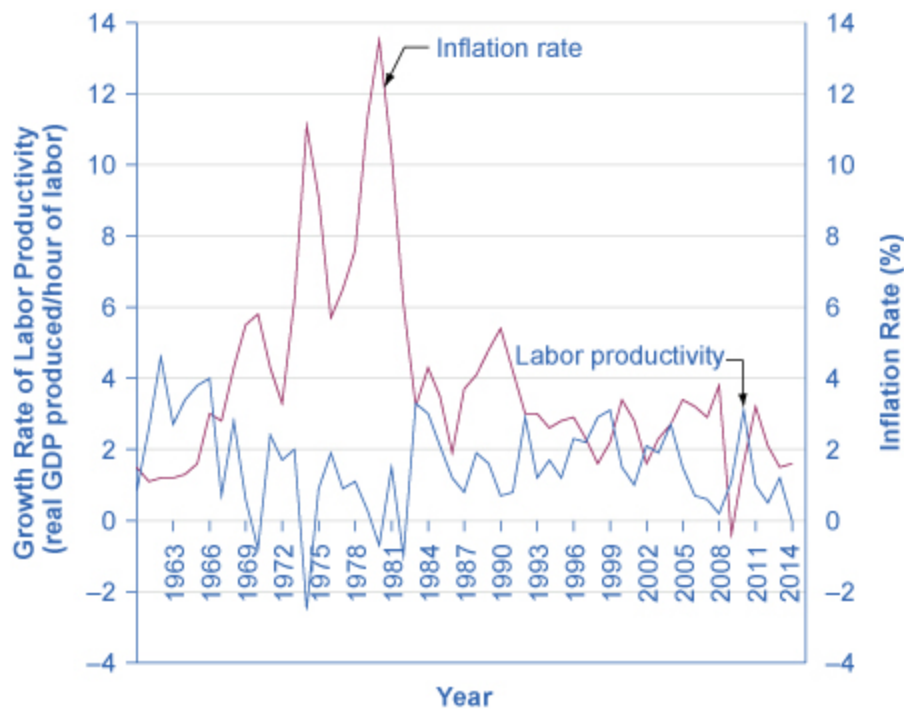
Inflation can make long-term planning difficult. In discussing unintended redistributions, we considered the case of someone trying to plan for retirement with a pension that is fixed in nominal terms and a high rate of inflation. Similar problems arise for all people trying to save for retirement, because they must consider what their money will really buy several

decades in the future when the rate of future inflation cannot be known with certainty.

Inflation, especially at moderate or high levels, will pose substantial planning problems for businesses, too. A firm can make money from inflation—for example, by paying bills and wages as late as possible so that it can pay in inflated dollars, while collecting revenues as soon as possible. A firm can also suffer losses from inflation, as in the case of a retail business that gets stuck holding too much cash, only to see the value of that cash eroded by inflation. But when a business spends its time focusing on how to profit by inflation, or at least how to avoid suffering from it, an inevitable tradeoff strikes: less time is spent on improving products and services or on figuring out how to make existing products and services more cheaply. An economy with high inflation rewards businesses that have found clever ways of profiting from inflation, which are not necessarily the businesses that excel at productivity, innovation, or quality of service.

In the short term, low or moderate levels of inflation may not pose an overwhelming difficulty for business planning, because costs of doing business and sales revenues may rise at similar rates. If, however, inflation varies substantially over the short or medium term, then it may make sense for businesses to stick to shorter-term strategies. The evidence as to whether relatively low rates of inflation reduce productivity is controversial among economists. There is some evidence that if inflation can be held to moderate levels of less than 3% per year, it need not prevent a nation's real economy from growing at a healthy pace. For some countries that have experienced hyperinflation of several thousand percent per year, an annual inflation rate of 20–30% may feel basically the same as zero. However, several economists have pointed to the suggestive fact that when U.S. inflation heated up in the early 1970s—to 10%—U.S. growth in productivity slowed down, and when inflation slowed down in the 1980s, productivity edged up again not long thereafter, as shown in [\[link\]](#).

U.S. Inflation Rate and U.S. Labor Productivity, 1961–2014



Over the last several decades in the United States, there have been times when rising inflation rates have been closely followed by lower productivity rates and lower inflation rates have corresponded to increasing productivity rates. As the graph shows, however, this correlation does not always exist.

Any Benefits of Inflation?

Although the economic effects of inflation are primarily negative, two countervailing points are worth noting. First, the impact of inflation will differ considerably according to whether it is creeping up slowly at 0% to 2% per year, galloping along at 10% to 20% per year, or racing to the point of hyperinflation at, say, 40% per month. Hyperinflation can rip an economy and a society apart. An annual inflation rate of 2%, 3%, or 4%, however, is a long way from a national crisis. Low inflation is also better than deflation which occurs with severe recessions.

Second, an argument is sometimes made that moderate inflation may help the economy by making wages in labor markets more flexible. The discussion in [Unemployment](#) pointed out that wages tend to be sticky in their downward movements and that unemployment can result. A little inflation could nibble away at real wages, and thus help real wages to decline if necessary. In this way, even if a moderate or high rate of inflation may act as sand in the gears of the economy, perhaps a low rate of inflation serves as oil for the gears of the labor market. This argument is controversial. A full analysis would have to take all the effects of inflation into account. It does, however, offer another reason to believe that, all things considered, very low rates of inflation may not be especially harmful.

Key Concepts and Summary

Unexpected inflation will tend to hurt those whose money received, in terms of wages and interest payments, does not rise with inflation. In contrast, inflation can help those who owe money that can be paid in less valuable, inflated dollars. Low rates of inflation have relatively little economic impact over the short term. Over the medium and the long term, even low rates of inflation can complicate future planning. High rates of inflation can muddle price signals in the short term and prevent market forces from operating efficiently, and can vastly complicate long-term savings and investment decisions.

Self-Check Question

Exercise:

Problem:

If inflation rises unexpectedly by 5%, would a state government that had recently borrowed money to pay for a new highway benefit or lose?

Solution:

The state government would benefit because it would repay the loan in less valuable dollars than it borrowed. Plus, tax revenues for the state government would increase because of the inflation.

Review Question

Exercise:

Problem:

Identify several parties likely to be helped and hurt by inflation.

Critical Thinking Questions

Exercise:

Problem:

If, over time, wages and salaries on average rise at least as fast as inflation, why do people worry about how inflation affects incomes?

Exercise:

Problem: Who in an economy is the big winner from inflation?

References

Shiller, Robert. "Why Do People Dislike Inflation?" *NBER Working Paper Series, National Bureau of Economic Research*, p. 52. 1996.

Indexing and Its Limitations

By the end of this section, you will be able to:

- Explain the relationship between indexing and inflation
- Identify three ways the government can control inflation through macroeconomic policy

When a price, wage, or interest rate is adjusted automatically with inflation, it is said to be **indexed**. An indexed payment increases according to the index number that measures inflation. A wide array of indexing arrangements is observed in private markets and government programs. Since the negative effects of inflation depend in large part on having inflation unexpectedly affect one part of the economy but not another—say, increasing the prices that people pay but not the wages that workers receive—indexing will take some of the sting out of inflation.

Indexing in Private Markets

In the 1970s and 1980s, labor unions commonly negotiated wage contracts that had **cost-of-living adjustments (COLAs)** which guaranteed that their wages would keep up with inflation. These contracts were sometimes written as, for example, COLA plus 3%. Thus, if inflation was 5%, the wage increase would automatically be 8%, but if inflation rose to 9%, the wage increase would automatically be 12%. COLAs are a form of indexing applied to wages.

Loans often have built-in inflation adjustments, too, so that if the inflation rate rises by two percentage points, then the interest rate charged on the loan rises by two percentage points as well. An **adjustable-rate mortgage (ARM)** is a kind of loan used to purchase a home in which the interest rate varies with the rate of inflation. Often, a borrower will be able receive a lower interest rate if borrowing with an ARM, compared to a fixed-rate loan. The reason is that with an ARM, the lender is protected against the risk that higher inflation will reduce the real loan payments, and so the risk premium part of the interest rate can be correspondingly lower.

A number of ongoing or long-term business contracts also have provisions that prices will be adjusted automatically according to inflation. Sellers like such contracts because they are not locked into a low nominal selling price if inflation turns out higher than expected; buyers like such contracts because they are not locked into a high buying price if inflation turns out to be lower than expected. A contract with automatic adjustments for inflation in effect agrees on a real price to be paid, rather than a nominal price.

Indexing in Government Programs

Many government programs are indexed to inflation. The U.S. income tax code is designed so that as a person's income rises above certain levels, the tax rate on the marginal income earned rises as well; this is what is meant by the expression "move into a higher tax bracket." For example, according to the basic tax tables from the Internal Revenue Service, in 2014 a single person owed 10% of all taxable income from \$0 to \$9,075; 15% of all income from \$9,076 to \$36,900; 25% of all taxable income from \$36,901 to \$89,350; 28% of all taxable income from \$89,351 to \$186,350; 33% of all taxable income from \$186,351 to \$405,100; 35% of all taxable income from \$405,101 to \$406,750; and 39.6% of all income from \$406,751 and above.

Because of the many complex provisions in the rest of the tax code, the taxes owed by any individual cannot be exactly determined based on these numbers, but the numbers illustrate the basic theme that tax rates rise as the marginal dollar of income rises. Until the late 1970s, if nominal wages increased along with inflation, people were moved into higher tax brackets and owed a higher proportion of their income in taxes, even though their real income had not risen. This "bracket creep," as it was called, was eliminated by law in 1981. Now, the income levels where higher tax rates kick in are indexed to rise automatically with inflation.

The Social Security program offers two examples of indexing. Since the passage of the Social Security Indexing Act of 1972, the level of Social Security benefits increases each year along with the Consumer Price Index. Also, Social Security is funded by payroll taxes, which are imposed on the income earned up to a certain amount—\$117,000 in 2014. This level of income is adjusted upward each year according to the rate of inflation, so

that the indexed rise in the benefit level is accompanied by an indexed increase in the Social Security tax base.

As yet another example of a government program affected by indexing, in 1996 the U.S. government began offering indexed bonds. Bonds are means by which the U.S. government (and many private-sector companies as well) borrows money; that is, investors buy the bonds, and then the government repays the money with interest. Traditionally, government bonds have paid a fixed rate of interest. This policy gave a government that had borrowed an incentive to encourage inflation, because it could then repay its past borrowing in inflated dollars at a lower real interest rate. But indexed bonds promise to pay a certain real rate of interest above whatever inflation rate occurs. In the case of a retiree trying to plan for the long term and worried about the risk of inflation, for example, indexed bonds that guarantee a rate of return higher than inflation—no matter the level of inflation—can be a very comforting investment.

Might Indexing Reduce Concern over Inflation?

Indexing may seem like an obviously useful step. After all, when individuals, firms, and government programs are indexed against inflation, then people can worry less about arbitrary redistributions and other effects of inflation.

However, some of the fiercest opponents of inflation express grave concern about indexing. They point out that indexing is always partial. Not every employer will provide COLAs for workers. Not all companies can assume that costs and revenues will rise in lockstep with the general rates of inflation. Not all interest rates for borrowers and savers will change to match inflation exactly. But as partial inflation indexing spreads, the political opposition to inflation may diminish. After all, older people whose Social Security benefits are protected against inflation, or banks that have loaned their money with adjustable-rate loans, no longer have as much reason to care whether inflation heats up. In a world where some people are indexed against inflation and some are not, financially savvy businesses and investors may seek out ways to be protected against inflation, while the financially unsophisticated and small businesses may suffer from it most.

A Preview of Policy Discussions of Inflation

This chapter has focused on how inflation is measured, historical experience with inflation, how to adjust nominal variables into real ones, how inflation affects the economy, and how indexing works. The causes of inflation have barely been hinted at, and government policies to deal with inflation have not been addressed at all. These issues will be taken up in depth in other chapters. However, it is useful to offer a preview here.

The cause of inflation can be summed up in one sentence: Too many dollars chasing too few goods. The great surges of inflation early in the twentieth century came after wars, which are a time when government spending is very high, but consumers have little to buy, because production is going to the war effort. Governments also commonly impose price controls during wartime. After the war, the price controls end and pent-up buying power surges forth, driving up inflation. On the other hand, if too few dollars are chasing too many goods, then inflation will decline or even turn into deflation. Therefore, slowdowns in economic activity, as in major recessions and the Great Depression, are typically associated with a reduction in inflation or even outright deflation.

The policy implications are clear. If inflation is to be avoided, the amount of purchasing power in the economy must grow at roughly the same rate as the production of goods. Macroeconomic policies that the government can use to affect the amount of purchasing power—through taxes, spending, and regulation of interest rates and credit—can thus cause inflation to rise or reduce inflation to lower levels.

Note:

A \$550 Million Loaf of Bread?

As we will learn in [Money and Banking](#), the existence of money provides enormous benefits to an economy. In a real sense, money is the lubrication that enhances the workings of markets. Money makes transactions easier. It allows people to find employment producing one product, then use the money earned to purchase the other products they need to live on. However, too much money in circulation can lead to inflation. Extreme

cases of governments recklessly printing money lead to hyperinflation. Inflation reduces the value of money. Hyperinflation, because money loses value so quickly, ultimately results in people no longer using money. The economy reverts to barter, or it adopts another country's more stable currency, like U.S. dollars. In the meantime, the economy literally falls apart as people leave jobs and fend for themselves because it is not worth the time to work for money that will be worthless in a few days. Only national governments have the power to cause hyperinflation. Hyperinflation typically happens when government faces extraordinary demands for spending, which it cannot finance by taxes or borrowing. The only option is to print money—more and more of it. With more money in circulation chasing the same amount (or even less) goods and services, the only result is higher and higher prices until the economy and/or the government collapses. This is why economists are generally wary of letting inflation get out of control.

Key Concepts and Summary

A payment is said to be indexed if it is automatically adjusted for inflation. Examples of indexing in the private sector include wage contracts with cost-of-living adjustments (COLAs) and loan agreements like adjustable-rate mortgages (ARMs). Examples of indexing in the public sector include tax brackets and Social Security payments.

Self-Check Questions

Exercise:

Problem:

How should an increase in inflation affect the interest rate on an adjustable-rate mortgage?

Solution:

Higher inflation reduces real interest rates on fixed rate mortgages. Because ARMs can be adjusted, higher inflation leads to higher interest rates on ARMs.

Exercise:

Problem:

A fixed-rate mortgage has the same interest rate over the life of the loan, whether the mortgage is for 15 or 30 years. By contrast, an adjustable-rate mortgage changes with market interest rates over the life of the mortgage. If inflation falls unexpectedly by 3%, what would likely happen to a homeowner with an adjustable-rate mortgage?

Solution:

Because the mortgage has an adjustable rate, the rate should fall by 3%, the same as inflation, to keep the real interest rate the same.

Review Questions

Exercise:

Problem: What is indexing?

Exercise:

Problem:

Name several forms of indexing in the private and public sector.

Critical Thinking Questions

Exercise:

Problem:

If a government gains from unexpected inflation when it borrows, why would it choose to offer indexed bonds?

Exercise:

Problem: Do you think perfect indexing is possible? Why or why not?

Problems

Exercise:

Problem:

If inflation rises unexpectedly by 5%, indicate for each of the following whether the economic actor is helped, hurt, or unaffected:

- a. A union member with a COLA wage contract
- b. Someone with a large stash of cash in a safe deposit box
- c. A bank lending money at a fixed rate of interest
- d. A person who is not due to receive a pay raise for another 11 months

Exercise:

Problem:

Rosalie the Retiree knows that when she retires in 16 years, her company will give her a one-time payment of \$20,000. However, if the inflation rate is 6% per year, how much buying power will that \$20,000 have when measured in today's dollars? *Hint:* Start by calculating the rise in the price level over the 16 years.

References

Wines, Michael. "How Bad is Inflation in Zimbabwe?" *The New York Times*, May 2, 2006.
http://www.nytimes.com/2006/05/02/world/africa/02zimbabwe.html?pagewanted=all&_r=0.

Hanke, Steve H. "R.I.P. Zimbabwe Dollar." *CATO Institute*. Accessed December 31, 2013. <http://www.cato.org/zimbabwe>.

Massachusetts Institute of Technology. 2015. "Billion Prices Project." Accessed March 4, 2015. <http://bpp.mit.edu/usa/>.

Glossary

adjustable-rate mortgage (ARM)

a loan used to purchase a home in which the interest rate varies with market interest rates

cost-of-living adjustments (COLAs)

a contractual provision that wage increases will keep up with inflation

indexed

a price, wage, or interest rate is adjusted automatically for inflation

Introduction to Government Budgets and Fiscal Policy

class="introduction"

Shut Downs and Parks

Yellowstone
National Park is
one of the many
national parks
forced to close
down during the
government shut
down in October
2013. (Credit:
modification of
work by
“daveynin”/flick
r Creative
Commons)



Note:**No Yellowstone Park?**

So you had trekked all the way to see Yellowstone National Park in the beautiful month of October 2013, only to find it... closed. Closed! Why? For two weeks in October 2013, the U.S. federal government shut down. Many federal services, like the national parks, closed and 800,000 federal employees were furloughed. Tourists were shocked and so was the rest of the world: Congress and the President could not agree on a budget. Inside the Capitol, Republicans and Democrats argued about spending priorities and whether to increase the national debt limit. Each year's budget, which is over \$3 trillion of spending, must be approved by Congress and signed by the President. Two thirds of the budget is entitlements and other mandatory spending which occur without congressional or presidential action once the programs are set up. Tied to the budget debate was the issue of increasing the debt ceiling—how high the national debt of the U.S. government can be. The House of Representatives refused to sign on to the bills to fund the government unless they included provisions to stop or change the Affordable Health Care Act (more colloquially known as Obamacare). As the days ticked by, the United States came very close to defaulting on its debt.

Why does the federal budget create such intense debates? What would happen if the United States actually defaulted on its debt? In this chapter, we will examine the federal budget, taxation, and fiscal policy. We will also look at the annual federal budget deficits and the national debt.

Note:**Introduction to Government Budgets and Fiscal Policy**

In this chapter, you will learn about:

- Government Spending
- Taxation
- Federal Deficits and the National Debt
- Using Fiscal Policy to Fight Recessions, Unemployment, and Inflation
- Automatic Stabilizers

- Practical Problems with Discretionary Fiscal Policy
- The Question of a Balanced Budget

All levels of government—federal, state, and local—have budgets that show how much revenue the government expects to receive in taxes and other income and how the government plans to spend it. Budgets, however, can shift dramatically within a few years, as policy decisions and unexpected events shake up earlier tax and spending plans.

In this chapter, we revisit fiscal policy, which was first covered in [Welcome to Economics!](#) Fiscal policy is one of two policy tools for fine tuning the economy (the other is monetary policy). While monetary policy is made by policymakers at the Federal Reserve, fiscal policy is made by Congress and the President.

The discussion of fiscal policy focuses on how federal government taxing and spending affects aggregate demand. All government spending and taxes affect the economy, but fiscal policy focuses strictly on the policies of the federal government. We begin with an overview of U.S. government spending and taxes. We then discuss fiscal policy from a short-run perspective; that is, how government uses tax and spending policies to address recession, unemployment, and inflation; how periods of recession and growth affect government budgets; and the merits of balanced budget proposals.

Government Spending

By the end of this section, you will be able to:

- Identify U.S. budget deficit and surplus trends over the past five decades
- Explain the differences between the U.S. federal budget, and state and local budgets

Government spending covers a range of services provided by the federal, state, and local governments. When the federal government spends more money than it receives in taxes in a given year, it runs a **budget deficit**. Conversely, when the government receives more money in taxes than it spends in a year, it runs a **budget surplus**. If government spending and taxes are equal, it is said to have a **balanced budget**. For example, in 2009, the U.S. government experienced its largest budget deficit ever, as the federal government spent \$1.4 trillion more than it collected in taxes. This deficit was about 10% of the size of the U.S. GDP in 2009, making it by far the largest budget deficit relative to GDP since the mammoth borrowing used to finance World War II.

This section presents an overview of government spending in the United States.

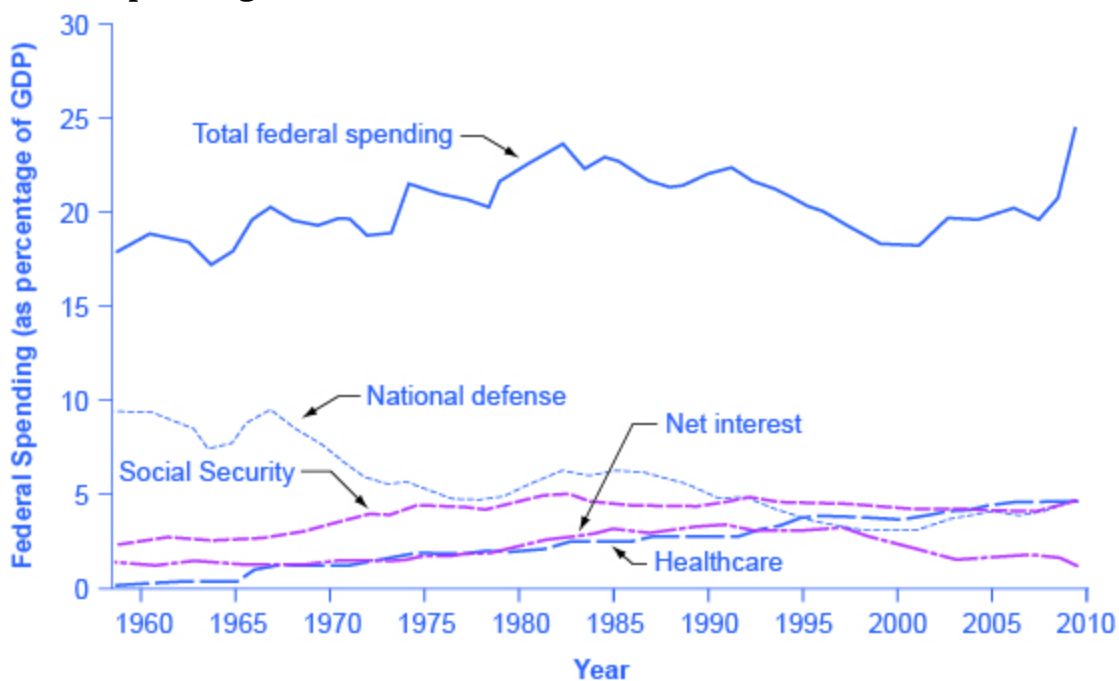
Total U.S. Government Spending

Federal spending in nominal dollars (that is, dollars not adjusted for inflation) has grown by a multiple of more than 38 over the last four decades, from \$93.4 billion in 1960 to \$3.9 trillion in 2014. Comparing spending over time in nominal dollars is misleading because it does not take into account inflation or growth in population and the real economy. A more useful method of comparison is to examine government spending as a percent of GDP over time.

The top line in [\[link\]](#) shows the level of federal spending since 1960, expressed as a share of GDP. Despite a widespread sense among many Americans that the federal government has been growing steadily larger, the

graph shows that federal spending has hovered in a range from 18% to 22% of GDP most of the time since 1960. The other lines in [\[link\]](#) show the major federal spending categories: national defense, Social Security, health programs, and interest payments. From the graph, we see that national defense spending as a share of GDP has generally declined since the 1960s, although there were some upward bumps in the 1980s buildup under President Ronald Reagan and in the aftermath of the terrorist attacks on September 11, 2001. In contrast, Social Security and healthcare have grown steadily as a percent of GDP. Healthcare expenditures include both payments for senior citizens (Medicare), and payments for low-income Americans (Medicaid). Medicaid is also partially funded by state governments. Interest payments are the final main category of government spending shown in the figure.

Federal Spending, 1960–2014



Since 1960, total federal spending has ranged from about 18% to 22% of GDP, although it climbed above that level in 2009, but quickly dropped back down to that level by 2013. The share spent on national defense has generally declined, while the share spent on Social Security and on healthcare expenses (mainly Medicare and Medicaid) has increased. (Source: *Economic Report of the*

President, Tables B-2 and B-22,
<http://www.gpo.gov/fdsys/pkg/ERP-2014/content-detail.html>)

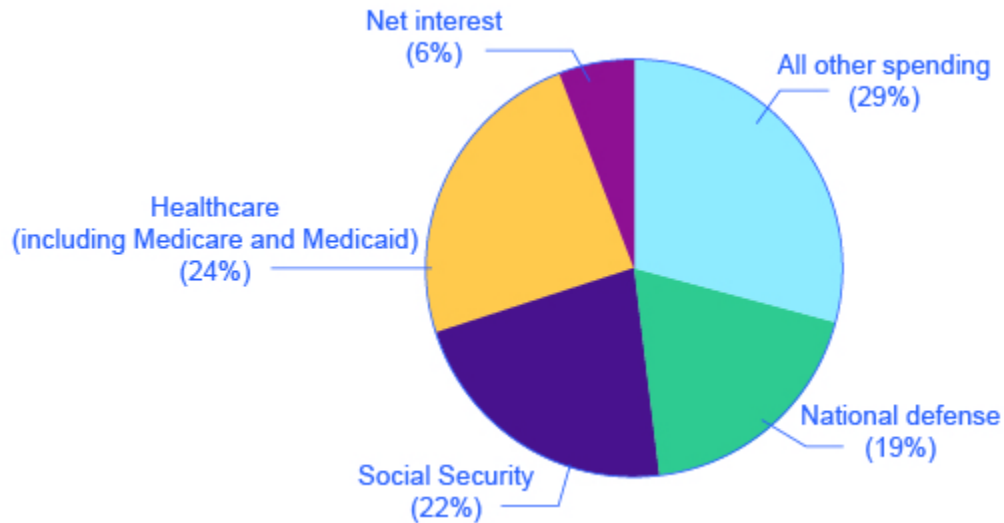
Each year, the government borrows funds from U.S. citizens and foreigners to cover its budget deficits. It does this by selling securities (Treasury bonds, notes, and bills)—in essence borrowing from the public and promising to repay with interest in the future. From 1961 to 1997, the U.S. government has run budget deficits, and thus borrowed funds, in almost every year. It had budget surpluses from 1998 to 2001, and then returned to deficits.

The interest payments on past federal government borrowing were typically 1–2% of GDP in the 1960s and 1970s but then climbed above 3% of GDP in the 1980s and stayed there until the late 1990s. The government was able to repay some of its past borrowing by running surpluses from 1998 to 2001 and, with help from low interest rates, the interest payments on past federal government borrowing had fallen back to 1.4% of GDP by 2012.

We investigate the patterns of government borrowing and debt in more detail later in this chapter, but first we need to clarify the difference between the deficit and the debt. *The deficit is not the debt.* The difference between the deficit and the debt lies in the time frame. The government deficit (or surplus) refers to what happens with the federal government budget each year. The government debt is accumulated over time; it is the sum of all past deficits and surpluses. If you borrow \$10,000 per year for each of the four years of college, you might say that your annual deficit was \$10,000, but your accumulated debt over the four years is \$40,000.

These four categories—national defense, Social Security, healthcare, and interest payments—account for roughly 73% of all federal spending, as [\[link\]](#) shows. The remaining 27% wedge of the pie chart covers all other categories of federal government spending: international affairs; science and technology; natural resources and the environment; transportation; housing; education; income support for the poor; community and regional development; law enforcement and the judicial system; and the administrative costs of running the government.

Slices of Federal Spending, 2014

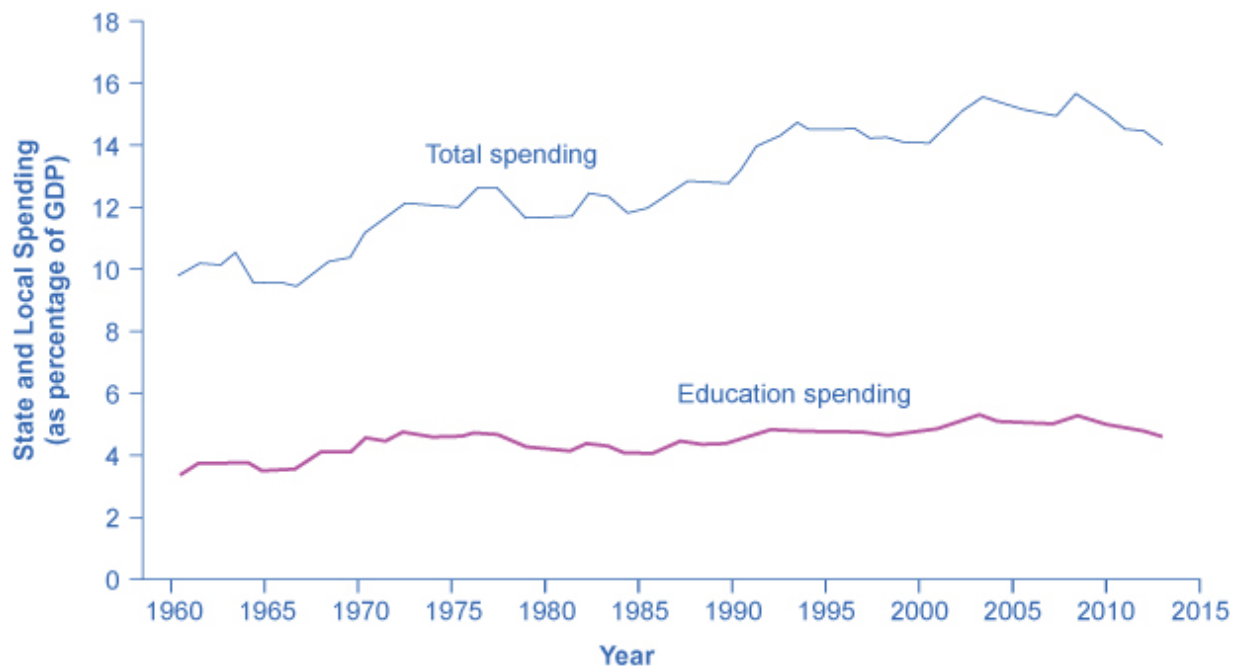


About 73% of government spending goes to four major areas: national defense, Social Security, healthcare, and interest payments on past borrowing. This leaves about 29% of federal spending for all other functions of the U.S. government. (Source: <https://www.whitehouse.gov/omb/budget/Historicals/>)

State and Local Government Spending

Although federal government spending often gets most of the media attention, state and local government spending is also substantial—at about \$3.1 trillion in 2014. [\[link\]](#) shows that state and local government spending has increased during the last four decades from around 8% to around 14% today. The single biggest item is education, which accounts for about one-third of the total. The rest covers programs like highways, libraries, hospitals and healthcare, parks, and police and fire protection. Unlike the federal government, all states (except Vermont) have balanced budget laws, which means any gaps between revenues and spending must be closed by higher taxes, lower spending, drawing down their previous savings, or some combination of all of these.

State and Local Spending, 1960–2013



Spending by state and local government increased from about 10% of GDP in the early 1960s to 14–16% by the mid-1970s. It has remained at roughly that level since. The single biggest spending item is education, including both K–12 spending and support for public colleges and universities, which has been about 4–5% of GDP in recent decades. Source: (Source: Bureau of Economic Analysis.)

U.S. presidential candidates often run for office pledging to improve the public schools or to get tough on crime. However, in the U.S. system of government, these tasks are primarily the responsibilities of state and local governments. Indeed, in fiscal year 2014 state and local governments spent about \$840 billion per year on education (including K–12 and college and university education), compared to only \$100 billion by the federal government, according to usgovernmentspending.com. In other words, about 90 cents of every dollar spent on education happens at the state and local level. A politician who really wants hands-on responsibility for reforming education or reducing crime might do better to run for mayor of a large city or for state governor rather than for president of the United States.

Key Concepts and Summary

Fiscal policy is the set of policies that relate to federal government spending, taxation, and borrowing. In recent decades, the level of federal government spending and taxes, expressed as a share of GDP, has not changed much, typically fluctuating between about 18% to 22% of GDP. However, the level of state spending and taxes, as a share of GDP, has risen from about 12–13% to about 20% of GDP over the last four decades. The four main areas of federal spending are national defense, Social Security, healthcare, and interest payments, which together account for about 70% of all federal spending. When a government spends more than it collects in taxes, it is said to have a budget deficit. When a government collects more in taxes than it spends, it is said to have a budget surplus. If government spending and taxes are equal, it is said to have a balanced budget. The sum of all past deficits and surpluses make up the government debt.

Self-Check Questions

Exercise:

Problem:

When governments run budget deficits, how do they make up the differences between tax revenue and spending?

Solution:

The government borrows funds by selling Treasury bonds, notes, and bills.

Exercise:

Problem:

When governments run budget surpluses, what is done with the extra funds?

Solution:

The funds can be used to pay down the national debt or else be refunded to the taxpayers.

Exercise:

Problem:

Is it possible for a nation to run budget deficits and still have its debt/GDP ratio fall? Explain your answer. Is it possible for a nation to run budget surpluses and still have its debt/GDP ratio rise? Explain your answer.

Solution:

Yes, a nation can run budget deficits and see its debt/GDP ratio fall. In fact, this is not uncommon. If the deficit is small in a given year, then the addition to debt in the numerator of the debt/GDP ratio will be relatively small, while the growth in GDP is larger, and so the debt/GDP ratio declines. This was the experience of the U.S. economy for the period from the end of World War II to about 1980. It is also theoretically possible, although not likely, for a nation to have a budget surplus and see its debt/GDP ratio rise. Imagine the case of a nation with a small surplus, but in a recession year when the economy shrinks. It is possible that the decline in the nation's debt, in the numerator of the debt/GDP ratio, would be proportionally less than the fall in the size of GDP, so the debt/GDP ratio would rise.

Review Questions

Exercise:

Problem:

Give some examples of changes in federal spending and taxes by the government that would be fiscal policy and some that would not.

Exercise:

Problem:

Have the spending and taxes of the U.S. federal government generally had an upward or a downward trend in the last few decades?

Exercise:

Problem:

What are the main categories of U.S. federal government spending?

Exercise:

Problem:

What is the difference between a budget deficit, a balanced budget, and a budget surplus?

Exercise:

Problem:

Have spending and taxes by state and local governments in the United States had a generally upward or downward trend in the last few decades?

Critical Thinking Questions

Exercise:

Problem:

Why is government spending typically measured as a percentage of GDP rather than in nominal dollars?

Exercise:

Problem:

Why are expenditures such as crime prevention and education typically done at the state and local level rather than at the federal level?

Exercise:

Problem:

Why is spending by the U.S. government on scientific research at NASA fiscal policy while spending by the University of Illinois is not fiscal policy? Why is a cut in the payroll tax fiscal policy whereas a cut in a state income tax is not fiscal policy?

Problems

Exercise:

Problem:

A government starts off with a total debt of \$3.5 billion. In year one, the government runs a deficit of \$400 million. In year two, the government runs a deficit of \$1 billion. In year three, the government runs a surplus of \$200 million. What is the total debt of the government at the end of year three?

References

Kramer, Mattea, et. al. *A People's Guide to the Federal Budget*. National Priorities Project. Northampton: Interlink Books, 2012.

Kurtzleben, Danielle. "10 States With The Largest Budget Shortfalls." *U.S. News & World Report*. January 14, 2011.
<http://www.usnews.com/news/articles/2011/01/14/10-states-with-the-largest-budget-shortfalls>.

Miller, Rich, and William Selway. "U.S. Cities and States Start Spending Again." *BloombergBusinessweek*, January 10, 2013.
<http://www.businessweek.com/articles/2013-01-10/u-dot-s-dot-cities-and-states-start-spending-again>.

Weisman, Jonathan. "After Year of Working Around Federal Cuts, Agencies Face Fewer Options." *The New York Times*, October 26, 2013.
http://www.nytimes.com/2013/10/27/us/politics/after-year-of-working-around-federal-cuts-agencies-face-fewer-options.html?_r=0.

Chantrill, Christopher. USGovernmentSpending.com. "Government Spending Details: United States Federal State and Local Government Spending, Fiscal Year 2013."
http://www.usgovernmentspending.com/year_spending_2013USbn_15bs2n_20.

Glossary

balanced budget

when government spending and taxes are equal

budget deficit

when the federal government spends more money than it receives in taxes in a given year

budget surplus

when the government receives more money in taxes than it spends in a year

Taxation

By the end of this section, you will be able to:

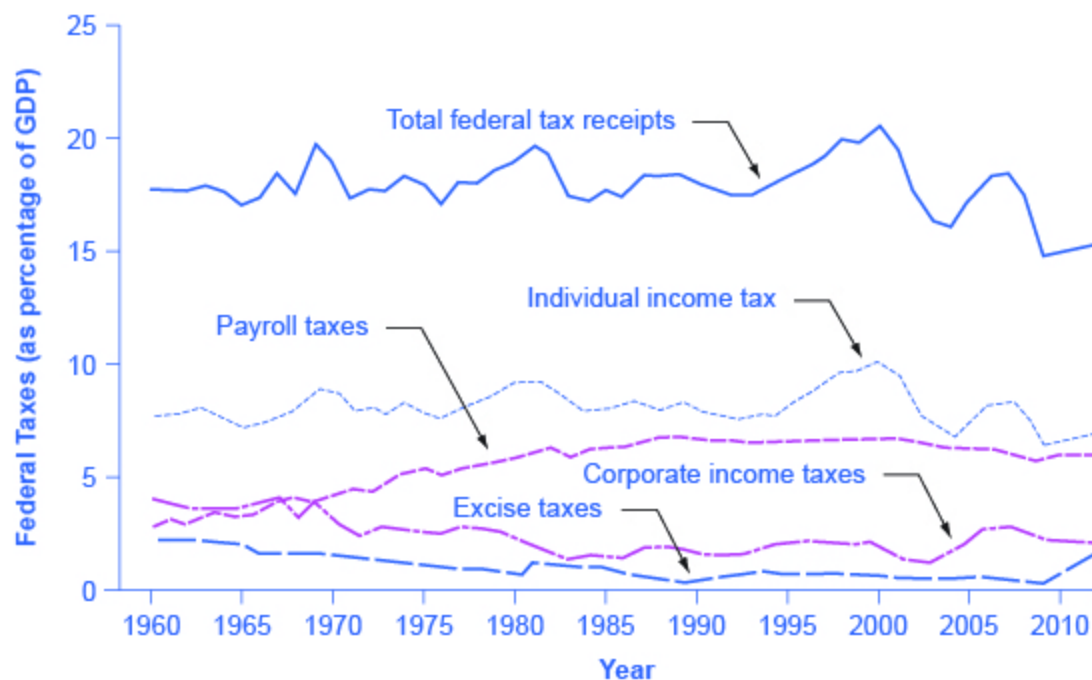
- Differentiate among a regressive tax, a proportional tax, and a progressive tax
- Identify the major sources of revenue for the U.S. federal budget

There are two main categories of taxes: those collected by the federal government and those collected by state and local governments. What percentage is collected and what that revenue is used for varies greatly. The following sections will briefly explain the taxation system in the United States.

Federal Taxes

Just as many Americans erroneously think that federal spending has grown considerably, many also believe that taxes have increased substantially. The top line of [\[link\]](#) shows total federal taxes as a share of GDP since 1960. Although the line rises and falls, it typically remains within the range of 17% to 20% of GDP, except for 2009, when taxes fell substantially below this level, due to recession.

Federal Taxes, 1960–2014



Federal tax revenues have been about 17–20% of GDP during most periods in recent decades. The primary sources of federal taxes are individual income taxes and the payroll taxes that finance Social Security and Medicare. Corporate income taxes and social insurance taxes provide smaller shares of revenue. (Source: *Economic Report of the President, 2015*. Table B-21, <https://www.whitehouse.gov/administration/eop/cea/economic-report-of-the-President/2015>)

[\[link\]](#) also shows the patterns of taxation for the main categories of taxes levied by the federal government: individual income taxes, corporate income taxes, and social insurance and retirement receipts. When most people think of taxes levied by the federal government, the first tax that comes to mind is the **individual income tax** that is due every year on April 15 (or the first business day after). The personal income tax is the largest single source of federal government revenue, but it still represents less than half of federal tax revenue.

The second largest source of federal revenue is the **payroll tax** (captured in social insurance and retirement receipts), which provides funds for Social Security and Medicare. Payroll taxes have increased steadily over time. Together, the personal income tax and the payroll tax accounted for about 80% of federal tax revenues in 2014. Although personal income tax revenues account for more total revenue than the payroll tax, nearly three-quarters of households pay more in payroll taxes than in income taxes.

The income tax is a **progressive tax**, which means that the tax rates increase as a household's income increases. Taxes also vary with marital status, family size, and other factors. The **marginal tax rates** (the tax that must be paid on all yearly income) for a single taxpayer range from 10% to 35%, depending on income, as the following Clear It Up feature explains.

Note:

How does the marginal rate work?

Suppose that a single taxpayer's income is \$35,000 per year. Also suppose that income from \$0 to \$9,075 is taxed at 10%, income from \$9,075 to \$36,900 is taxed at 15%, and, finally, income from \$36,900 and beyond is taxed at 25%. Since this person earns \$35,000, their marginal tax rate is 15%.

The key fact here is that the federal income tax is designed so that tax rates increase as income increases, up to a certain level. The payroll taxes that support Social Security and Medicare are designed in a different way. First, the payroll taxes for Social Security are imposed at a rate of 12.4% up to a certain wage limit, set at \$118,500 in 2015. Medicare, on the other hand, pays for elderly healthcare, and is fixed at 2.9%, with no upper ceiling.

In both cases, the employer and the employee split the payroll taxes. An employee only sees 6.2% deducted from his paycheck for Social Security, and 1.45% from Medicare. However, as economists are quick to point out, the employer's half of the taxes are probably passed along to the employees

in the form of lower wages, so in reality, the worker pays all of the payroll taxes.

The Medicare payroll tax is also called a **proportional tax**; that is, a flat percentage of all wages earned. The Social Security payroll tax is proportional up to the wage limit, but above that level it becomes a **regressive tax**, meaning that people with higher incomes pay a smaller share of their income in tax.

The third-largest source of federal tax revenue, as shown in [\[link\]](#) is the **corporate income tax**. The common name for corporate income is “profits.” Over time, corporate income tax receipts have declined as a share of GDP, from about 4% in the 1960s to an average of 1% to 2% of GDP in the first decade of the 2000s.

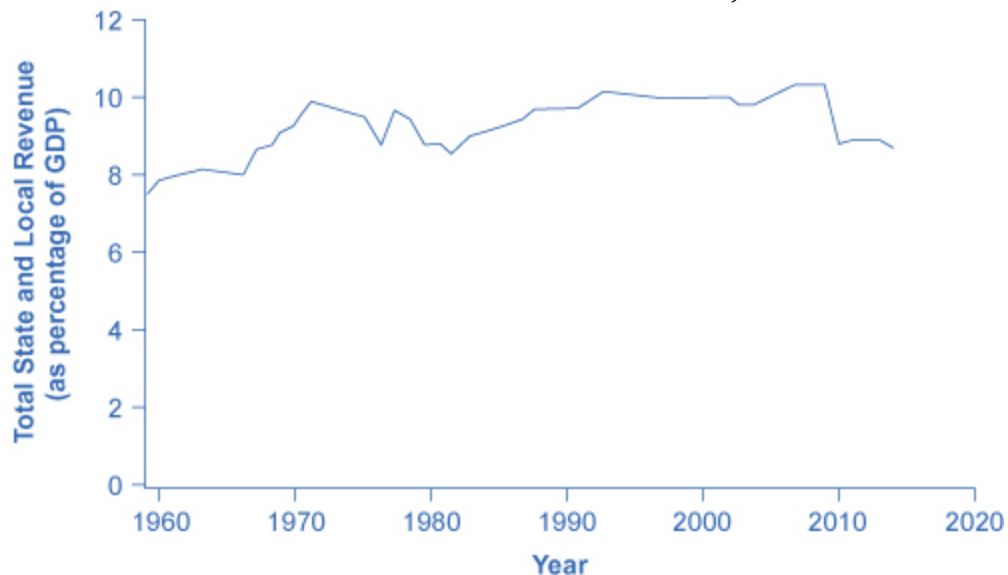
The federal government has a few other, smaller sources of revenue. It imposes an **excise tax**—that is, a tax on a particular good—on gasoline, tobacco, and alcohol. As a share of GDP, the amount collected by these taxes has stayed nearly constant over time, from about 2% of GDP in the 1960s to roughly 3% by 2014, according to the nonpartisan Congressional Budget Office. The government also imposes an **estate and gift tax** on people who pass large amounts of assets to the next generation—either after death or during life in the form of gifts. These estate and gift taxes collected about 0.2% of GDP in the first decade of the 2000s. By a quirk of legislation, the estate and gift tax was repealed in 2010, but reinstated in 2011. Other federal taxes, which are also relatively small in magnitude, include tariffs collected on imported goods and charges for inspections of goods entering the country.

State and Local Taxes

At the state and local level, taxes have been rising as a share of GDP over the last few decades to match the gradual rise in spending, as [\[link\]](#) illustrates. The main revenue sources for state and local governments are sales taxes, property taxes, and revenue passed along from the federal government, but many state and local governments also levy personal and corporate income taxes, as well as impose a wide variety of fees and

charges. The specific sources of tax revenue vary widely across state and local governments. Some states rely more on property taxes, some on sales taxes, some on income taxes, and some more on revenues from the federal government.

State and Local Tax Revenue as a Share of GDP, 1960–2014



State and local tax revenues have increased to match the rise in state and local spending. (Source: *Economic Report of the President, 2015*. Table B-21, <https://www.whitehouse.gov/administration/eop/cea/economic-report-of-the-President/2015>)

Key Concepts and Summary

The two main federal taxes are individual income taxes and payroll taxes that provide funds for Social Security and Medicare; these taxes together account for more than 80% of federal revenues. Other federal taxes include the corporate income tax, excise taxes on alcohol, gasoline and tobacco, and the estate and gift tax. A progressive tax is one, like the federal income tax, where those with higher incomes pay a higher share of taxes out of their income than those with lower incomes. A proportional tax is one, like the payroll tax for Medicare, where everyone pays the same share of taxes

regardless of income level. A regressive tax is one, like the payroll tax (above a certain threshold) that supports Social Security, where those with high income pay a lower share of income in taxes than those with lower incomes.

Self-Check Questions

Exercise:

Problem:

Suppose that gifts were taxed at a rate of 10% for amounts up to \$100,000 and 20% for anything over that amount. Would this tax be regressive or progressive?

Solution:

Progressive. People who give larger gifts subject to the higher tax rate would typically have larger incomes as well.

Exercise:

Problem:

If an individual owns a corporation for which he is the only employee, which different types of federal tax will he have to pay?

Solution:

Corporate income tax on his profits, individual income tax on his salary, and payroll tax taken out of the wages he pays himself.

Exercise:

Problem:

What taxes would an individual pay if he were self-employed and the business is not incorporated?

Solution:

individual income taxes

Exercise:

Problem:

The social security tax is 6.2% on employees' income earned below \$113,000. Is this tax progressive, regressive or proportional?

Solution:

The tax is regressive because wealthy income earners are not taxed at all on income above \$113,000. As a percent of total income, the social security tax hits lower income earners harder than wealthier individuals.

Review Questions

Exercise:

Problem:

What are the main categories of U.S. federal government taxes?

Exercise:

Problem:

What is the difference between a progressive tax, a proportional tax, and a regressive tax?

Critical Thinking Questions

Exercise:

Problem:

Excise taxes on tobacco and alcohol and state sales taxes are often criticized for being regressive. Although everyone pays the same rate regardless of income, why might this be so?

Exercise:**Problem:**

What is the benefit of having state and local taxes on income instead of collecting all such taxes at the federal level?

References

Burman, Leonard E., and Joel Selmrod. *Taxes in America: What Everyone Needs to Know*. New York: Oxford University Press, 2012.

Hall, Robert E., and Alvin Rabushka. *The Flat Tax (Hoover Classics)*. Stanford: Hoover Institution Press, 2007.

Kliff, Sarah. “How Congress Paid for Obamacare (in Two Charts).” *The Washington Post: WonkBlog* (blog), August 30, 2012.
<http://www.washingtonpost.com/blogs/wonkblog/wp/2012/08/30/how-congress-paid-for-obamacare-in-two-charts/>.

Matthews, Dylan. “America’s Taxes are the Most Progressive in the World. Its Government is Among the Least.” *The Washington Post: WonkBlog* (blog). April 5, 2013.
<http://www.washingtonpost.com/blogs/wonkblog/wp/2013/04/05/americas-taxes-are-the-most-progressive-in-the-world-its-government-is-among-the-least/>.

Glossary

corporate income tax

a tax imposed on corporate profits

estate and gift tax

a tax on people who pass assets to the next generation—either after death or during life in the form of gifts

excise tax

a tax on a specific good—on gasoline, tobacco, and alcohol

individual income tax

a tax based on the income, of all forms, received by individuals

marginal tax rates

or the tax that must be paid on all yearly income

payroll tax

a tax based on the pay received from employers; the taxes provide funds for Social Security and Medicare

progressive tax

a tax that collects a greater share of income from those with high incomes than from those with lower incomes

proportional tax

a tax that is a flat percentage of income earned, regardless of level of income

regressive tax

a tax in which people with higher incomes pay a smaller share of their income in tax

Federal Deficits and the National Debt

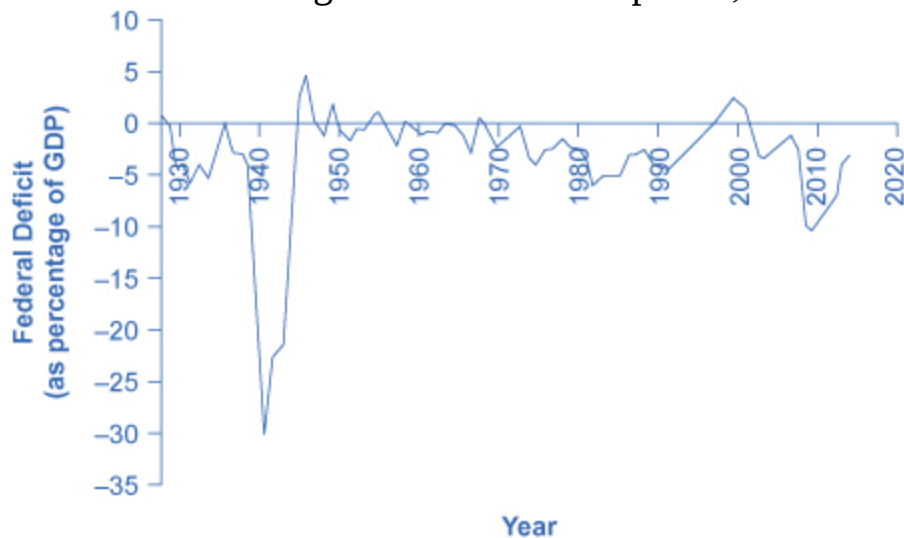
By the end of this section, you will be able to:

- Explain the U.S. federal budget in terms of annual debt and accumulated debt
- Understand how economic growth or decline can influence a budget surplus or budget deficit

Having discussed the revenue (taxes) and expense (spending) side of the budget, we now turn to the annual budget deficit or surplus, which is the difference between the tax revenue collected and spending over a fiscal year, which starts October 1 and ends September 30 of the next year.

[\[link\]](#) shows the pattern of annual federal budget deficits and surpluses, back to 1930, as a share of GDP. When the line is above the horizontal axis, the budget is in surplus; when the line is below the horizontal axis, a budget deficit occurred. Clearly, the biggest deficits as a share of GDP during this time were incurred to finance World War II. Deficits were also large during the 1930s, the 1980s, the early 1990s, and most recently during the recession of 2008–2009.

Pattern of Federal Budget Deficits and Surpluses, 1929–2014



The federal government has run budget deficits for decades. The budget was briefly in surplus in the late 1990s, before heading into deficit again in the first decade

of the 2000s—and especially deep deficits in the recession of 2008–2009. (Source: Federal Reserve Bank of St. Louis (FRED)).

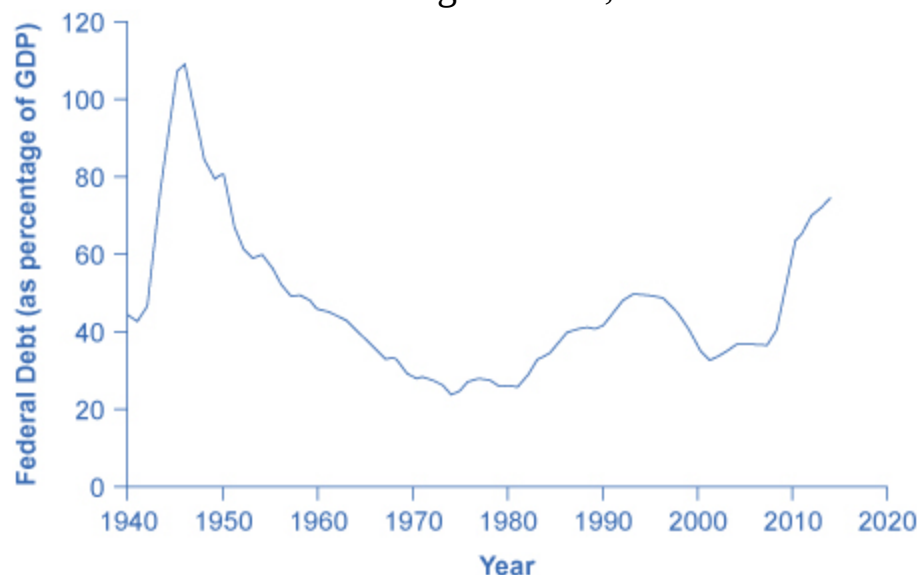
<http://research.stlouisfed.org/fred2/series/FYFSGDA188S>

)

Debt/GDP Ratio

Another useful way to view the budget deficit is through the prism of accumulated debt rather than annual deficits. The **national debt** refers to the total amount that the government has borrowed over time; in contrast, the budget deficit refers to how much has been borrowed in one particular year. [\[link\]](#) shows the ratio of debt/GDP since 1940. Until the 1970s, the debt/GDP ratio revealed a fairly clear pattern of federal borrowing. The government ran up large deficits and raised the debt/GDP ratio in World War II, but from the 1950s to the 1970s the government ran either surpluses or relatively small deficits, and so the debt/GDP ratio drifted down. Large deficits in the 1980s and early 1990s caused the ratio to rise sharply. When budget surpluses arrived from 1998 to 2001, the debt/GDP ratio declined substantially. The budget deficits starting in 2002 then tugged the debt/GDP ratio higher—with a big jump when the recession took hold in 2008–2009.

Federal Debt as a Percentage of GDP, 1942–2014



Federal debt is the sum of annual budget deficits and surpluses. Annual deficits do not always mean that the debt/GDP ratio is rising. During the 1960s and 1970s, the government often ran small deficits, but since the debt was growing more slowly than the economy, the debt/GDP ratio was declining over this time. In the 2008–2009 recession, the debt/GDP ratio rose sharply. (Source: *Economic Report of the President, Table B-20*, <http://www.gpo.gov/fdsys/pkg/ERP-2015/content-detail.html>)

The next Clear it Up feature discusses how the government handles the national debt.

Note:

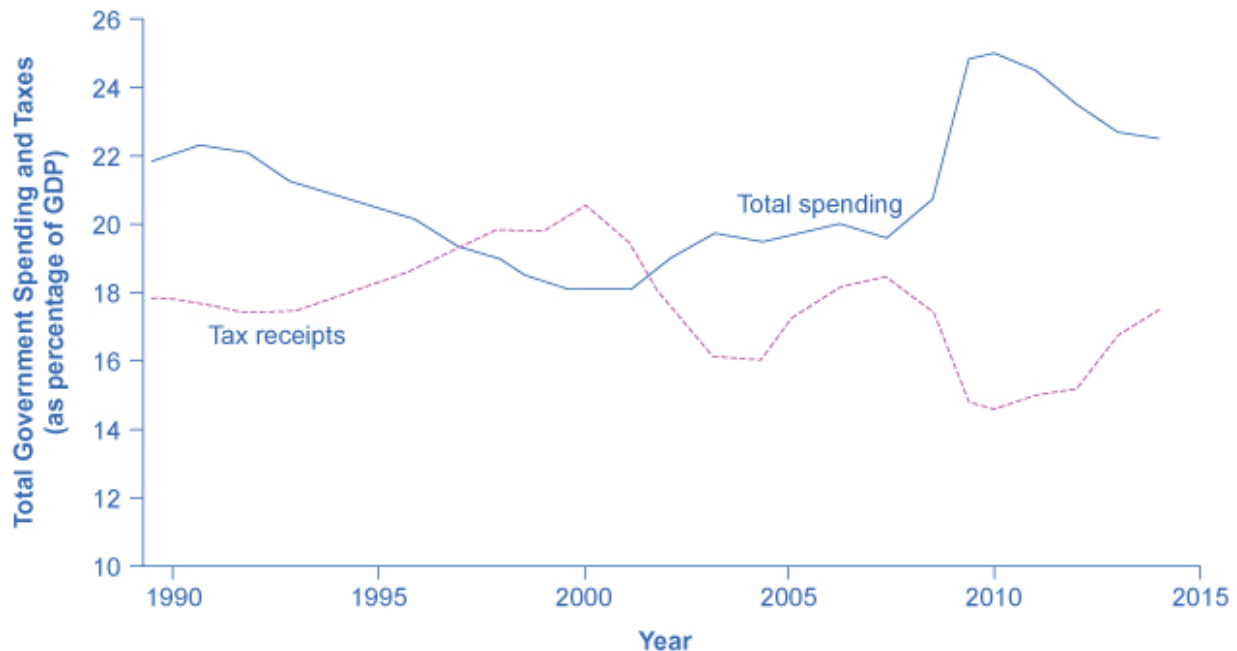
What is the national debt?

One year's federal budget deficit causes the federal government to sell Treasury bonds to make up the difference between spending programs and tax revenues. The dollar value of all the outstanding Treasury bonds on which the federal government owes money is equal to the national debt.

The Path from Deficits to Surpluses to Deficits

Why did the budget deficits suddenly turn to surpluses from 1998 to 2001? And why did the surpluses return to deficits in 2002? Why did the deficit become so large after 2007? [\[link\]](#) suggests some answers. The graph combines the earlier information on total federal spending and taxes in a single graph, but focuses on the federal budget since 1990.

Total Government Spending and Taxes as a Share of GDP, 1990–2014



When government spending exceeds taxes, the gap is the budget deficit. When taxes exceed spending, the gap is a budget surplus. The recessionary period starting in late 2007 saw higher spending and lower taxes, combining to create a large deficit in 2009. (Source: *Economic Report of the President, Tables B-21 and B-1*, "<http://www.gpo.gov/fdsys/pkg/ERP-2015/content-detail.html>")

Government spending as a share of GDP declined steadily through the 1990s. The biggest single reason was that defense spending declined from 5.2% of GDP in 1990 to 3.0% in 2000, but interest payments by the federal government also fell by about 1.0% of GDP. However, federal tax collections increased substantially in the later 1990s, jumping from 18.1% of GDP in 1994 to 20.8% in 2000. Powerful economic growth in the late 1990s fueled the boom in taxes. Personal income taxes rise as income goes up; payroll taxes rise as jobs and payrolls go up; corporate income taxes rise as profits go up. At the same time, government spending on transfer payments such as unemployment benefits, foods stamps, and welfare declined with more people working.

This sharp increase in tax revenues and decrease in expenditures on transfer payments was largely unexpected even by experienced budget analysts, and so budget surpluses came as a surprise. But in the early 2000s, many of these factors started running in reverse. Tax revenues sagged, due largely to the recession that started in March 2001, which reduced revenues. A series of tax cuts was enacted by Congress and signed into law by President George W. Bush, starting in 2001. In addition, government spending swelled due to increases in defense, healthcare, education, Social Security, and support programs for those who were hurt by the recession and the slow growth that followed. Deficits returned. When the severe recession hit in late 2007, spending climbed and tax collections fell to historically unusual levels, resulting in enormous deficits.

Longer-term forecasts of the U.S. budget, a decade or more into the future, predict enormous deficits. The higher deficits run during the recession of 2008–2009 have repercussions, and the demographics will be challenging. The primary reason is the “baby boom”—the exceptionally high birthrates that began in 1946, right after World War II, and lasted for about two decades. Starting in 2010, the front edge of the baby boom generation began to reach age 65, and in the next two decades, the proportion of Americans over the age of 65 will increase substantially. The current level of the payroll taxes that support Social Security and Medicare will fall well short of the projected expenses of these programs, as the following Clear It Up feature shows; thus, the forecast is for large budget deficits. A decision to collect more revenue to support these programs or to decrease benefit levels would alter this long-term forecast.

Note:

What is the long-term budget outlook for Social Security and Medicare? In 1946, just one American in 13 was over age 65. By 2000, it was one in eight. By 2030, one American in five will be over age 65. Two enormous U.S. federal programs focus on the elderly—Social Security and Medicare. The growing numbers of elderly Americans will increase spending on these programs, as well as on Medicaid. The current payroll tax levied on workers, which supports all of Social Security and the hospitalization

insurance part of Medicare, will not be enough to cover the expected costs. So, what are the options?

Long-term projections from the Congressional Budget Office in 2009 are that Medicare and Social Security spending combined will rise from 8.3% of GDP in 2009 to about 13% by 2035 and about 20% in 2080. If this rise in spending occurs, without any corresponding rise in tax collections, then some mix of changes must occur: (1) taxes will need to be increased dramatically; (2) other spending will need to be cut dramatically; (3) the retirement age and/or age receiving Medicare benefits will need to increase, or (4) the federal government will need to run extremely large budget deficits.

Some proposals suggest removing the cap on wages subject to the payroll tax, so that those with very high incomes would have to pay the tax on the entire amount of their wages. Other proposals suggest moving Social Security and Medicare from systems in which workers pay for retirees toward programs that set up accounts where workers save funds over their lifetimes and then draw out after retirement to pay for healthcare.

The United States is not alone in this problem. Indeed, providing the promised level of retirement and health benefits to a growing proportion of elderly with a falling proportion of workers is an even more severe problem in many European nations and in Japan. How to pay promised levels of benefits to the elderly will be a difficult public policy decision.

In the next module we shift to the use of fiscal policy to counteract business cycle fluctuations. In addition, we will explore proposals requiring a balanced budget—that is, for government spending and taxes to be equal each year. [The Impacts of Government Borrowing](#) will also cover how fiscal policy and government borrowing will affect national saving—and thus affect economic growth and trade imbalances.

Key Concepts and Summary

For most of the twentieth century, the U.S. government took on debt during wartime and then paid down that debt slowly in peacetime. However, it took on quite substantial debts in peacetime in the 1980s and early 1990s,

before a brief period of budget surpluses from 1998 to 2001, followed by a return to annual budget deficits since 2002, with very large deficits in the recession of 2008 and 2009. A budget deficit or budget surplus is measured annually. Total government debt or national debt is the sum of budget deficits and budget surpluses over time.

Self-Check Questions

Exercise:

Problem:

Debt has a certain self-reinforcing quality to it. There is one category of government spending that automatically increases along with the federal debt. What is it?

Solution:

As debt increases, interest payments also rise, so that the deficit grows even if we keep other government spending constant.

Exercise:

Problem: True or False:

- a. Federal spending has grown substantially in recent decades.
- b. By world standards, the U.S. government controls a relatively large share of the U.S. economy.
- c. A majority of the federal government's revenue is collected through personal income taxes.
- d. Education spending is slightly larger at the federal level than at the state and local level.
- e. State and local government spending has not risen much in recent decades.
- f. Defense spending is higher now than ever.
- g. The share of the economy going to federal taxes has increased substantially over time.

- h. Foreign aid is a large portion, although less than half, of federal spending.
 - i. Federal deficits have been very large for the last two decades.
 - j. The accumulated federal debt as a share of GDP is near an all-time high.
-

Solution:

- a. As a share of GDP, this is false. In nominal dollars, it is true.
- b. False.
- c. False.
- d. False. Education spending is much higher at the state level.
- e. False. As a share of GDP, it is up about 50.
- f. As a share of GDP, this is false, and in real dollars, it is also false.
- g. False.
- h. False; it's about 1%.
- i. False. Although budget deficits were large in 2003 and 2004, and continued into the later 2000s, the federal government ran budget surpluses from 1998–2001.
- j. False.

Review Questions

Exercise:

Problem:

What has been the general pattern of U.S. budget deficits in recent decades?

Exercise:

Problem:

What is the difference between a budget deficit and the national debt?

Critical Thinking Questions

Exercise:

Problem:

In a booming economy, is the federal government more likely to run surpluses or deficits? What are the various factors at play?

Exercise:

Problem:

Economist Arthur Laffer famously pointed out that, in some cases, income tax revenue can actually go up when tax rates go down. Why might this be the case?

Exercise:

Problem:

Is it possible for a nation to run budget deficits and still have its debt/GDP ratio fall? Explain your answer. Is it possible for a nation to run budget surpluses and still have its debt/GDP ratio rise? Explain your answer.

Problems

Exercise:

Problem:

If a government runs a budget deficit of \$10 billion dollars each year for ten years, then a surplus of \$1 billion for five years, and then a balanced budget for another ten years, what is the government debt?

References

Eisner, Robert. *The Great Deficit Scares: The Federal Budget, Trade, and Social Security*. New York: Priority Press Publications, 1997.

Weisman, Jonathan, and Ashley Parker. "Republicans Back Down, Ending Crisis Over Shutdown and Debt Limit." *The New York Times*, October 16, 2013. <http://www.nytimes.com/2013/10/17/us/congress-budget-debate.html>.

Wessel, David. *Red Ink: Inside the High-Stakes Politics of Federal Budget*. New York: Crown Publishing Group, 2013.

Glossary

national debt

the total accumulated amount the government has borrowed, over time, and not yet paid back

Using Fiscal Policy to Fight Recession, Unemployment, and Inflation

By the end of this section, you will be able to:

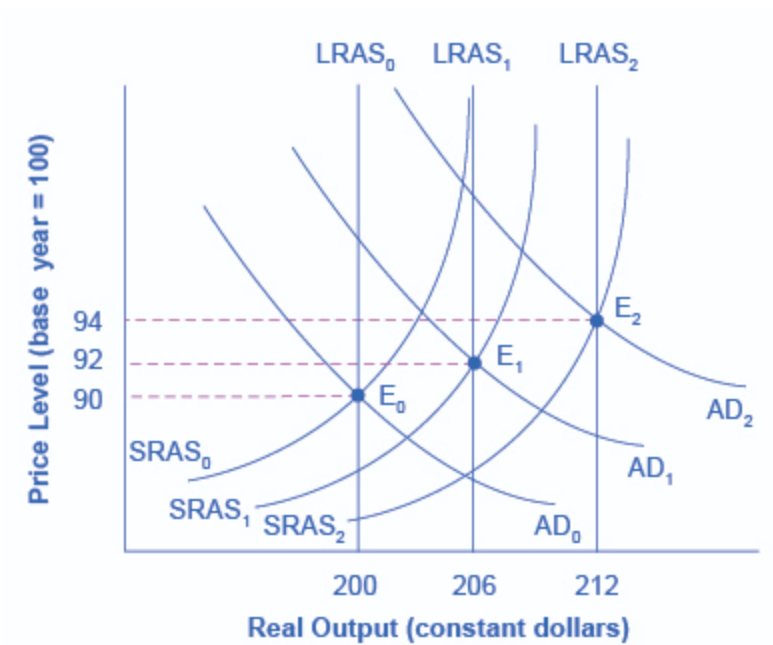
- Explain how expansionary fiscal policy can shift aggregate demand and influence the economy
- Explain how contractionary fiscal policy can shift aggregate demand and influence the economy

We need to emphasize that fiscal policy is the use of government spending and tax policy to alter the economy. Fiscal policy does not include all spending (such as the increase in spending that accompanies a war).

Graphically, we see that fiscal policy, whether through change in spending or taxes, shifts the aggregate demand outward in the case of **expansionary fiscal policy** and inward in the case of **contractionary fiscal policy**. [\[link\]](#) illustrates the process by using an aggregate demand/aggregate supply diagram in a growing economy. The original equilibrium occurs at E_0 , the intersection of aggregate demand curve AD_0 and aggregate supply curve $SRAS_0$, at an output level of 200 and a price level of 90.

One year later, aggregate supply has shifted to the right to $SRAS_1$ in the process of long-term economic growth, and aggregate demand has also shifted to the right to AD_1 , keeping the economy operating at the new level of potential GDP. The new equilibrium (E_1) is an output level of 206 and a price level of 92. One more year later, aggregate supply has again shifted to the right, now to $SRAS_2$, and aggregate demand shifts right as well to AD_2 . Now the equilibrium is E_2 , with an output level of 212 and a price level of 94. In short, the figure shows an economy that is growing steadily year to year, producing at its potential GDP each year, with only small inflationary increases in the price level.

A Healthy, Growing Economy



In this well-functioning economy, each year aggregate supply and aggregate demand shift to the right so that the economy proceeds from equilibrium E_0 to E_1 to E_2 . Each year, the economy produces at potential GDP with only a small inflationary increase in the price level. But if aggregate demand does not smoothly shift to the right and match increases in aggregate supply, growth with deflation can develop.

Aggregate demand and aggregate supply do not always move neatly together. Aggregate demand may fail to increase along with aggregate supply, or aggregate demand may even shift left, for a number of possible reasons: households become hesitant about consuming; firms decide against investing as much; or perhaps the demand from other countries for exports diminishes. For example, investment by private firms in physical capital in the U.S. economy boomed during the late 1990s, rising from 14.1% of GDP in 1993 to 17.2% in 2000, before falling back to 15.2% by 2002.

Conversely, if shifts in aggregate demand run ahead of increases in aggregate supply, inflationary increases in the price level will result. Business cycles of recession and recovery are the consequence of shifts in aggregate supply and aggregate demand.

[Monetary Policy and Bank Regulation](#) shows us that a central bank can use its powers over the banking system to engage in countercyclical—or “against the business cycle”—actions. If recession threatens, the central bank uses an expansionary monetary policy to increase the supply of money, increase the quantity of loans, reduce interest rates, and shift aggregate demand to the right. If inflation threatens, the central bank uses contractionary monetary policy to reduce the supply of money, reduce the quantity of loans, raise interest rates, and shift aggregate demand to the left. Fiscal policy is another macroeconomic policy tool for adjusting aggregate demand by using either government spending or taxation policy.

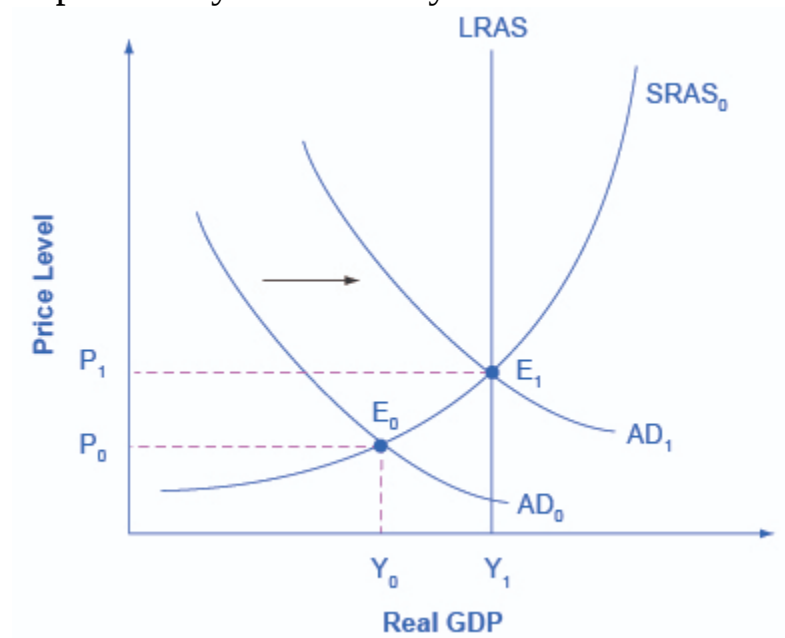
Expansionary Fiscal Policy

Expansionary fiscal policy increases the level of aggregate demand, through either increases in government spending or reductions in taxes. Expansionary policy can do this by (1) increasing consumption by raising disposable income through cuts in personal income taxes or payroll taxes; (2) increasing investments by raising after-tax profits through cuts in business taxes; and (3) increasing government purchases through increased spending by the federal government on final goods and services and raising federal grants to state and local governments to increase their expenditures on final goods and services. Contractionary fiscal policy does the reverse: it decreases the level of aggregate demand by decreasing consumption, decreasing investments, and decreasing government spending, either through cuts in government spending or increases in taxes. The aggregate demand/aggregate supply model is useful in judging whether expansionary or contractionary fiscal policy is appropriate.

Consider first the situation in [\[link\]](#), which is similar to the U.S. economy during the recession in 2008–2009. The intersection of aggregate demand (AD_0) and aggregate supply ($SRAS_0$) is occurring below the level of potential GDP as indicated by the LRAS curve. At the equilibrium (E_0), a

recession occurs and unemployment rises. In this case, expansionary fiscal policy using tax cuts or increases in government spending can shift aggregate demand to AD_1 , closer to the full-employment level of output. In addition, the price level would rise back to the level P_1 associated with potential GDP.

Expansionary Fiscal Policy



The original equilibrium (E_0) represents a recession, occurring at a quantity of output (Y_0) below potential GDP. However, a shift of aggregate demand from AD_0 to AD_1 , enacted through an expansionary fiscal policy, can move the economy to a new equilibrium output of E_1 at the level of potential GDP which is shown by the LRAS curve. Since the economy was originally producing below potential GDP, any inflationary increase in the price level from P_0 to P_1 that results should be relatively small.

Should the government use tax cuts or spending increases, or a mix of the two, to carry out expansionary fiscal policy? After the Great Recession of 2008–2009 (which started, actually, in very late 2007), U.S. government spending rose from 19.6% of GDP in 2007 to 24.6% in 2009, while tax revenues declined from 18.5% of GDP in 2007 to 14.8% in 2009. The choice between whether to use tax or spending tools often has a political tinge. As a general statement, conservatives and Republicans prefer to see expansionary fiscal policy carried out by tax cuts, while liberals and Democrats prefer that expansionary fiscal policy be implemented through spending increases. The Obama administration and Congress passed an \$830 billion expansionary policy in early 2009 involving both tax cuts and increases in government spending, according to the Congressional Budget Office. However, state and local governments, whose budgets were also hard hit by the recession, began cutting their spending—a policy that offset federal expansionary policy.

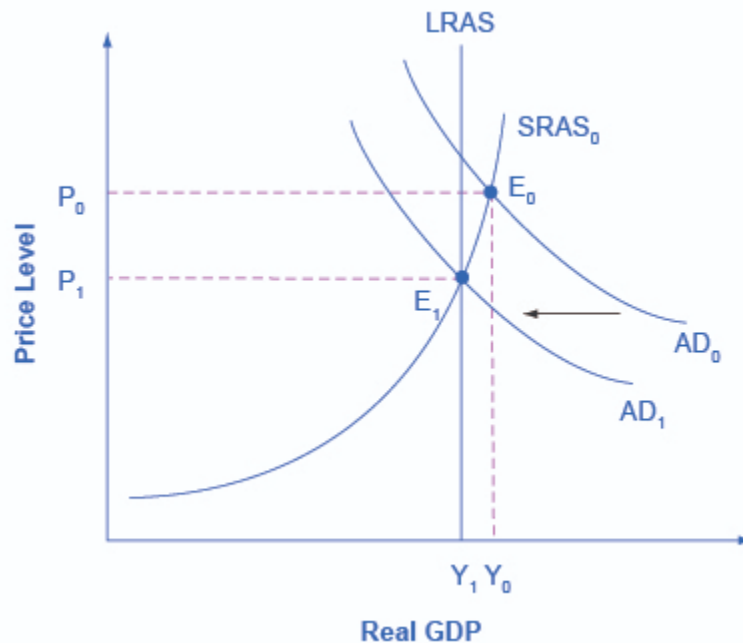
The conflict over which policy tool to use can be frustrating to those who want to categorize economics as “liberal” or “conservative,” or who want to use economic models to argue against their political opponents. But the AD–AS model can be used both by advocates of smaller government, who seek to reduce taxes and government spending, and by advocates of bigger government, who seek to raise taxes and government spending. Economic studies of specific taxing and spending programs can help to inform decisions about whether taxes or spending should be changed, and in what ways. Ultimately, decisions about whether to use tax or spending mechanisms to implement macroeconomic policy is, in part, a political decision rather than a purely economic one.

Contractionary Fiscal Policy

Fiscal policy can also contribute to pushing aggregate demand beyond potential GDP in a way that leads to inflation. As shown in [\[link\]](#), a very large budget deficit pushes up aggregate demand, so that the intersection of aggregate demand (AD_0) and aggregate supply ($SRAS_0$) occurs at equilibrium E_0 , which is an output level above potential GDP. This is sometimes known as an “overheating economy” where demand is so high that there is upward pressure on wages and prices, causing inflation. In this

situation, contractionary fiscal policy involving federal spending cuts or tax increases can help to reduce the upward pressure on the price level by shifting aggregate demand to the left, to AD_1 , and causing the new equilibrium E_1 to be at potential GDP, where aggregate demand intersects the LRAS curve.

A Contractionary Fiscal Policy



The economy starts at the equilibrium quantity of output Y_0 , which is above potential GDP. The extremely high level of aggregate demand will generate inflationary increases in the price level. A contractionary fiscal policy can shift aggregate demand down from AD_0 to AD_1 , leading to a new equilibrium output E_1 , which occurs at potential GDP, where AD_1 intersects the LRAS curve.

Again, the AD-AS model does not dictate how this contractionary fiscal policy is to be carried out. Some may prefer spending cuts; others may prefer tax increases; still others may say that it depends on the specific situation. The model only argues that, in this situation, aggregate demand needs to be reduced.

Key Concepts and Summary

Expansionary fiscal policy increases the level of aggregate demand, either through increases in government spending or through reductions in taxes. Expansionary fiscal policy is most appropriate when an economy is in recession and producing below its potential GDP. Contractionary fiscal policy decreases the level of aggregate demand, either through cuts in government spending or increases in taxes. Contractionary fiscal policy is most appropriate when an economy is producing above its potential GDP.

Self-Check Questions

Exercise:

Problem:

What is the main reason for employing contractionary fiscal policy in a time of strong economic growth?

Solution:

To keep prices from rising too much or too rapidly.

Exercise:

Problem:

What is the main reason for employing expansionary fiscal policy during a recession?

Solution:

To increase employment.

Review Questions

Exercise:

Problem:

What is the difference between expansionary fiscal policy and contractionary fiscal policy?

Exercise:**Problem:**

Under what general macroeconomic circumstances might a government use expansionary fiscal policy? When might it use contractionary fiscal policy?

Critical Thinking Questions**Exercise:****Problem:**

How will cuts in state budget spending affect federal expansionary policy?

Exercise:**Problem:**

Is expansionary fiscal policy more attractive to politicians who believe in larger government or to politicians who believe in smaller government? Explain your answer.

Problems**Exercise:**

Problem:

Specify whether expansionary or contractionary fiscal policy would seem to be most appropriate in response to each of the situations below and sketch a diagram using aggregate demand and aggregate supply curves to illustrate your answer:

- a. A recession.
- b. A stock market collapse that hurts consumer and business confidence.
- c. Extremely rapid growth of exports.
- d. Rising inflation.
- e. A rise in the natural rate of unemployment.
- f. A rise in oil prices.

References

Alesina, Alberto, and Francesco Giavazzi. *Fiscal Policy after the Financial Crisis (National Bureau of Economic Research Conference Report)*. Chicago: University Of Chicago Press, 2013.

Martin, Fernando M. "Fiscal Policy in the Great Recession and Lessons from the Past." Federal Reserve Bank of St. Louis: *Economic Synopses*. no. 1 (2012). http://research.stlouisfed.org/publications/es/12/ES_2012-01-06.pdf.

Bivens, Josh, Andrew Fieldhouse, and Heidi Shierholz. "From Free-fall to Stagnation: Five Years After the Start of the Great Recession, Extraordinary Policy Measures Are Still Needed, But Are Not Forthcoming." *Economic Policy Institute*. Last modified February 14, 2013. <http://www.epi.org/publication/bp355-five-years-after-start-of-great-recession/>.

Lucking, Brian, and Dan Wilson. Federal Reserve Bank of San Francisco, "FRBSF Economic Letter—U.S. Fiscal Policy: Headwind or Tailwind?"

Last modified July 2, 2012. <http://www.frbsf.org/economic-research/publications/economic-letter/2012/july/us-fiscal-policy/>.

Greenstone, Michael, and Adam Looney. Brookings. “The Role of Fiscal Stimulus in the Ongoing Recovery.” Last modified July 6, 2012. <http://www.brookings.edu/blogs/jobs/posts/2012/07/06-jobs-greenstone-looney>.

Glossary

contractionary fiscal policy

fiscal policy that decreases the level of aggregate demand, either through cuts in government spending or increases in taxes

expansionary fiscal policy

fiscal policy that increases the level of aggregate demand, either through increases in government spending or cuts in taxes

Automatic Stabilizers

By the end of this section, you will be able to:

- Describe how discretionary fiscal policy can be used by the federal government to stabilize the economy.
- Identify examples of automatic stabilizers.
- Understand how a standardized employment budget can be used to identify automatic stabilizers.

The millions of unemployed in 2008–2009 could collect unemployment insurance benefits to replace some of their salaries. Federal fiscal policies include **discretionary fiscal policy**, when the government passes a new law that explicitly changes tax or spending levels. The stimulus package of 2009 is an example. Changes in tax and spending levels can also occur automatically, due to **automatic stabilizers**, such as unemployment insurance and food stamps, which are programs that are already laws that stimulate aggregate demand in a recession and hold down aggregate demand in a potentially inflationary boom.

Counterbalancing Recession and Boom

Consider first the situation where aggregate demand has risen sharply, causing the equilibrium to occur at a level of output above potential GDP. This situation will increase inflationary pressure in the economy. The policy prescription in this setting would be a dose of contractionary fiscal policy, implemented through some combination of higher taxes and lower spending. To some extent, *both* changes happen automatically. On the tax side, a rise in aggregate demand means that workers and firms throughout the economy earn more. Because taxes are based on personal income and corporate profits, a rise in aggregate demand automatically increases tax payments. On the spending side, stronger aggregate demand typically means lower unemployment and fewer layoffs, and so there is less need for government spending on unemployment benefits, welfare, Medicaid, and other programs in the social safety net.

The process works in reverse, too. If aggregate demand were to fall sharply so that a recession occurs, then the prescription would be for expansionary fiscal policy—some mix of tax cuts and spending increases. The lower level of aggregate demand and higher unemployment will tend to pull down personal incomes and corporate profits, an effect that will reduce the amount of taxes owed automatically. Higher unemployment and a weaker economy should lead to increased government spending on unemployment benefits, welfare, and other similar domestic programs. In 2009, the stimulus package included an extension in the time allowed to collect unemployment insurance. In addition, the automatic stabilizers react to a weakening of aggregate demand with expansionary fiscal policy and react to a strengthening of aggregate demand with contractionary fiscal policy, just as the AD/AS analysis suggests.

The very large budget deficit of 2009 was produced by a combination of automatic stabilizers and discretionary fiscal policy. The Great Recession, starting in late 2007, meant less tax-generating economic activity, which triggered the automatic stabilizers that reduce taxes. Most economists, even those who are concerned about a possible pattern of persistently large budget deficits, are much less concerned or even quite supportive of larger budget deficits in the short run of a few years during and immediately after a severe recession.

A glance back at economic history provides a second illustration of the power of automatic stabilizers. Remember that the length of economic upswings between recessions has become longer in the U.S. economy in recent decades (as discussed in [Unemployment](#)). The three longest economic booms of the twentieth century happened in the 1960s, the 1980s, and the 1991–2001 time period. One reason why the economy has tipped into recession less frequently in recent decades is that the size of government spending and taxes has increased in the second half of the twentieth century. Thus, the automatic stabilizing effects from spending and taxes are now larger than they were in the first half of the twentieth century. Around 1900, for example, federal spending was only about 2% of GDP. In 1929, just before the Great Depression hit, government spending was still just 4% of GDP. In those earlier times, the smaller size of government made

automatic stabilizers far less powerful than in the last few decades, when government spending often hovers at 20% of GDP or more.

The Standardized Employment Deficit or Surplus

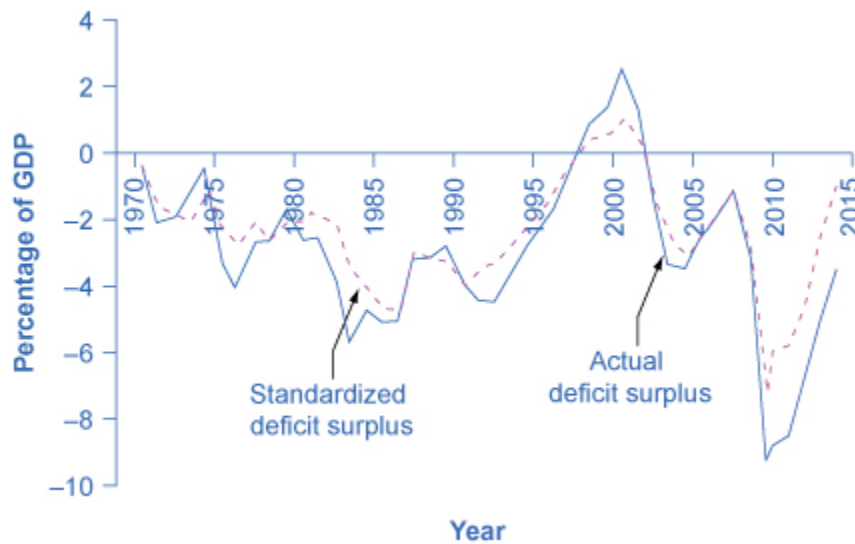
Each year, the nonpartisan Congressional Budget Office (CBO) calculates the **standardized employment budget**—that is, what the budget deficit or surplus would be if the economy were producing at potential GDP, where people who look for work were finding jobs in a reasonable period of time and businesses were making normal profits, with the result that both workers and businesses would be earning more and paying more taxes. In effect, the standardized employment deficit eliminates the impact of the automatic stabilizers. [\[link\]](#) compares the actual budget deficits of recent decades with the CBO's standardized deficit.

Note:

Visit this [website](#) to learn more from the Congressional Budget Office.



Comparison of Actual Budget Deficits with the Standardized Employment Deficit



When the economy is in recession, the standardized employment budget deficit is less than the actual budget deficit because the economy is below potential GDP, and the automatic stabilizers are reducing taxes and increasing spending. When the economy is performing extremely well, the standardized employment deficit (or surplus) is higher than the actual budget deficit (or surplus) because the economy is producing about potential GDP, so the automatic stabilizers are increasing taxes and reducing the need for government spending.

(Sources: *Actual and Cyclically Adjusted Budget Surpluses/Deficits*,

<http://www.cbo.gov/publication/43977>; and *Economic Report of the President, Table B-1*, <http://www.gpo.gov/fdsys/pkg/ERP-2013/content-detail.html>)

Notice that in recession years, like the early 1990s, 2001, or 2009, the standardized employment deficit is smaller than the actual deficit. During recessions, the automatic stabilizers tend to increase the budget deficit, so if

the economy was instead at full employment, the deficit would be reduced. However, in the late 1990s the standardized employment budget surplus was lower than the actual budget surplus. The gap between the standardized budget deficit or surplus and the actual budget deficit or surplus shows the impact of the automatic stabilizers. More generally, the standardized budget figures allow you to see what the budget deficit would look like with the economy held constant—at its potential GDP level of output.

Automatic stabilizers occur quickly. Lower wages means that a lower amount of taxes is withheld from paychecks right away. Higher unemployment or poverty means that government spending in those areas rises as quickly as people apply for benefits. However, while the automatic stabilizers offset part of the shifts in aggregate demand, they do not offset all or even most of it. Historically, automatic stabilizers on the tax and spending side offset about 10% of any initial movement in the level of output. This offset may not seem enormous, but it is still useful. Automatic stabilizers, like shock absorbers in a car, can be useful if they reduce the impact of the worst bumps, even if they do not eliminate the bumps altogether.

Key Concepts and Summary

Fiscal policy is conducted both through discretionary fiscal policy, which occurs when the government enacts taxation or spending changes in response to economic events, or through automatic stabilizers, which are taxing and spending mechanisms that, by their design, shift in response to economic events without any further legislation. The standardized employment budget is the calculation of what the budget deficit or budget surplus would have been in a given year if the economy had been producing at its potential GDP in that year. Many economists and politicians criticize the use of fiscal policy for a variety of reasons, including concerns over time lags, the impact on interest rates, and the inherently political nature of fiscal policy. We cover the critique of fiscal policy in the next module.

Self-Check Questions

Exercise:

Problem:

In a recession, does the actual budget surplus or deficit fall above or below the standardized employment budget?

Solution:

It falls below because less tax revenue than expected is collected.

Exercise:**Problem:**

What is the main advantage of automatic stabilizers over discretionary fiscal policy?

Solution:

Automatic stabilizers take effect very quickly, whereas discretionary policy can take a long time to implement.

Exercise:**Problem:**

Explain how automatic stabilizers work, both on the taxation side and on the spending side, first in a situation where the economy is producing less than potential GDP and then in a situation where the economy is producing more than potential GDP.

Solution:

In a recession, because of the decline in economic output, less income is earned, and so less in taxes is automatically collected. Many welfare and unemployment programs are designed so that those who fall into certain categories, like “unemployed” or “low income,” are eligible for benefits. During a recession, more people fall into these categories and become eligible for benefits automatically. The combination of reduced taxes and higher spending is just what is needed for an economy in recession producing below potential GDP. With an

economic boom, average income levels rise in the economy, so more in taxes is automatically collected. Fewer people meet the criteria for receiving government assistance to the unemployed or the needy, so government spending on unemployment assistance and welfare falls automatically. This combination of higher taxes and lower spending is just what is needed if an economy is producing above its potential GDP.

Review Questions

Exercise:

Problem:

What is the difference between discretionary fiscal policy and automatic stabilizers?

Exercise:

Problem: Why do automatic stabilizers function “automatically?”

Exercise:

Problem: What is the standardized employment budget?

Critical Thinking Questions

Exercise:

Problem:

Is Medicaid (federal government aid to low-income families and individuals) an automatic stabilizer?

Glossary

automatic stabilizers

tax and spending rules that have the effect of slowing down the rate of decrease in aggregate demand when the economy slows down and restraining aggregate demand when the economy speeds up, without any additional change in legislation

discretionary fiscal policy

the government passes a new law that explicitly changes overall tax or spending levels with the intent of influencing the level or overall economic activity

standardized employment budget

the budget deficit or surplus in any given year adjusted for what it would have been if the economy were producing at potential GDP

Practical Problems with Discretionary Fiscal Policy

By the end of this section, you will be able to:

- Understand how fiscal policy and monetary policy are interconnected
- Explain the three lag times that often occur when solving economic problems.
- Identify the legal and political challenges of responding to an economic problem.

In the early 1960s, many leading economists believed that the problem of the business cycle, and the swings between cyclical unemployment and inflation, were a thing of the past. On the cover of its December 31, 1965, issue, *Time* magazine, then the premier news magazine in the United States, ran a picture of John Maynard Keynes, and the story inside identified Keynesian theories as “the prime influence on the world’s economies.” The article reported that policymakers have “used Keynesian principles not only to avoid the violent [business] cycles of prewar days but to produce phenomenal economic growth and to achieve remarkably stable prices.”

This happy consensus, however, did not last. The U.S. economy suffered one recession from December 1969 to November 1970, a deeper recession from November 1973 to March 1975, and then double-dip recessions from January to June 1980 and from July 1981 to November 1982. At various times, inflation and unemployment both soared. Clearly, the problems of macroeconomic policy had not been completely solved. As economists began to consider what had gone wrong, they identified a number of issues that make discretionary fiscal policy more difficult than it had seemed in the rosy optimism of the mid-1960s.

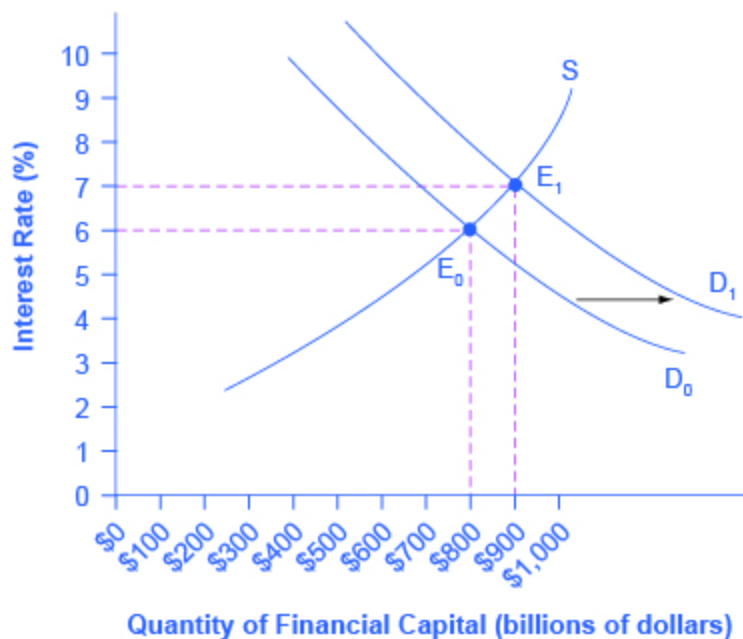
Fiscal Policy and Interest Rates

Because fiscal policy affects the quantity that the government borrows in financial capital markets, it not only affects aggregate demand—it can also affect interest rates. In [\[link\]](#), the original equilibrium (E_0) in the financial capital market occurs at a quantity of \$800 billion and an interest rate of 6%. However, an increase in government budget deficits shifts the demand

for financial capital from D_0 to D_1 . The new equilibrium (E_1) occurs at a quantity of \$900 billion and an interest rate of 7%.

A consensus estimate based on a number of studies is that an increase in budget deficits (or a fall in budget surplus) by 1% of GDP will cause an increase of 0.5–1.0% in the long-term interest rate.

Fiscal Policy and Interest Rates



When a government borrows money in the financial capital market, it causes a shift in the demand for financial capital from D_0 to D_1 . As the equilibrium moves from E_0 to E_1 , the equilibrium interest rate rises from 6% to 7% in this example. In this way, an expansionary fiscal policy intended to shift aggregate demand to the right can also lead to a higher interest rate, which has the effect of shifting aggregate demand back to the left.

A problem arises here. An expansionary fiscal policy, with tax cuts or spending increases, is intended to increase aggregate demand. If an expansionary fiscal policy also causes higher interest rates, then firms and households are discouraged from borrowing and spending (as occurs with tight monetary policy), thus reducing aggregate demand. Even if the direct effect of expansionary fiscal policy on increasing demand is not totally offset by lower aggregate demand from higher interest rates, fiscal policy can end up being less powerful than was originally expected. This is referred to as **crowding out**, where government borrowing and spending results in higher interest rates, which reduces business investment and household consumption.

The broader lesson is that fiscal and monetary policy must be coordinated. If expansionary fiscal policy is to work well, then the central bank can also reduce or keep short-term interest rates low. Conversely, monetary policy can also help to ensure that contractionary fiscal policy does not lead to a recession.

Long and Variable Time Lags

Monetary policy can be changed several times each year, but fiscal policy is much slower to be enacted. Imagine that the economy starts to slow down. It often takes some months before the economic statistics signal clearly that a downturn has started, and a few months more to confirm that it is truly a recession and not just a one- or two-month blip. The time it takes to determine that a recession has occurred is often called the **recognition lag**. After this lag, policymakers become aware of the problem and propose fiscal policy bills. The bills go into various congressional committees for hearings, negotiations, votes, and then, if passed, eventually for the president's signature. Many fiscal policy bills about spending or taxes propose changes that would start in the next budget year or would be phased in gradually over time. The time to get a bill passed is often referred to as the **legislative lag**. Finally, once the bill is passed it takes some time for the funds to be dispersed to the appropriate agencies to implement the programs. The time to get the projects started is often called the **implementation lag**.

Moreover, the exact level of fiscal policy to be implemented is never completely clear. Should the budget deficit be increased by 0.5% of GDP? By 1% of GDP? By 2% of GDP? In an AD/AS diagram, it is straightforward to sketch an aggregate demand curve shifting to the potential GDP level of output. In the real world, the actual level of potential output is known only roughly, not precisely, and exactly how a spending cut or tax increase will affect aggregate demand is always somewhat controversial. Also unknown is the state of the economy at any point in time. During the early days of the Obama administration, for example, no one knew how deep in the hole the economy really was. During the financial crisis of 2008-09, the rapid collapse of the banking system and automotive sector made it difficult to assess how quickly the economy was collapsing.

Thus, it can take many months or even more than a year to begin an expansionary fiscal policy after a recession has started—and even then, uncertainty will remain over exactly how much to expand or contract taxes and spending. When politicians attempt to use countercyclical fiscal policy to fight recession or inflation, they run the risk of responding to the macroeconomic situation of two or three years ago, in a way that may be exactly wrong for the economy at that time. George P. Schultz, a professor of economics, former Secretary of the Treasury, and Director of the Office of Management and Budget, once wrote: “While the economist is accustomed to the concept of lags, the politician likes instant results. The tension comes because, as I have seen on many occasions, the economist’s lag is the politician’s nightmare.”

Temporary and Permanent Fiscal Policy

A temporary tax cut or spending increase will explicitly last only for a year or two, and then revert back to its original level. A permanent tax cut or spending increase is expected to stay in place for the foreseeable future. The effect of temporary and permanent fiscal policies on aggregate demand can be very different. Consider how you would react if the government announced a tax cut that would last one year and then be repealed, in comparison with how you would react if the government announced a

permanent tax cut. Most people and firms will react more strongly to a permanent policy change than a temporary one.

This fact creates an unavoidable difficulty for countercyclical fiscal policy. The appropriate policy may be to have an expansionary fiscal policy with large budget deficits during a recession, and then a contractionary fiscal policy with budget surpluses when the economy is growing well. But if both policies are explicitly temporary ones, they will have a less powerful effect than a permanent policy.

Structural Economic Change Takes Time

When an economy recovers from a recession, it does not usually revert back to its exact earlier shape. Instead, the internal structure of the economy evolves and changes and this process can take time. For example, much of the economic growth of the mid-2000s was in the sectors of construction (especially of housing) and finance. However, when housing prices started falling in 2007 and the resulting financial crunch led into recession (as discussed in [Monetary Policy and Bank Regulation](#)), both sectors contracted. The manufacturing sector of the U.S. economy has been losing jobs in recent years as well, under pressure from technological change and foreign competition. Many of the people thrown out of work from these sectors in the Great Recession of 2008–2009 will never return to the same jobs in the same sectors of the economy; instead, the economy will need to grow in new and different directions, as the following Clear It Up feature shows. Fiscal policy can increase overall demand, but the process of structural economic change—the expansion of a new set of industries and the movement of workers to those industries—inevitably takes time.

Note:

Why do jobs vanish?

People can lose jobs for a variety of reasons: because of a recession, but also because of longer-run changes in the economy, such as new technology. Productivity improvements in auto manufacturing, for example, can reduce the number of workers needed, and eliminate these

jobs in the long run. The Internet has created jobs but also caused the loss of jobs as well, from travel agents to book store clerks. Many of these jobs may never come back. Short-run fiscal policy to reduce unemployment can create jobs, but it cannot replace jobs that will never return.

The Limitations of Fiscal Policy

Fiscal policy can help an economy that is producing below its potential GDP to expand aggregate demand so that it produces closer to potential GDP, thus lowering unemployment. But fiscal policy cannot help an economy produce at an output level above potential GDP without causing inflation. At this point, unemployment becomes so low that workers become scarce and wages rise rapidly.

Note:

Visit this [website](#) to read about how the recovery is being affected by fiscal policies.



Political Realities and Discretionary Fiscal Policy

A final problem for discretionary fiscal policy arises out of the difficulties of explaining to politicians how countercyclical fiscal policy that runs against the tide of the business cycle should work. Politicians often have a gut-level belief that when the economy and tax revenues slow down, it is

time to hunker down, pinch pennies, and trim expenses. Countercyclical policy, however, says that when the economy has slowed down, it is time for the government to go on a spree, raising spending, and cutting taxes. This offsets the drop in the economy in the other sectors. Conversely, when economic times are good and tax revenues are rolling in, politicians often feel that it is time for tax cuts and new spending. But countercyclical policy says that this economic boom should be an appropriate time for keeping taxes high and restraining spending.

Politicians tend to prefer expansionary fiscal policy over contractionary policy. There is rarely a shortage of proposals for tax cuts and spending increases, especially during recessions. However, politicians are less willing to hear the message that in good economic times, they should propose tax increases and spending limits. In the economic upswing of the late 1990s and early 2000s, for example, the U.S. GDP grew rapidly. Estimates from respected government economic forecasters like the nonpartisan Congressional Budget Office and the Office of Management and Budget stated that the GDP was above potential GDP, and that unemployment rates were unsustainably low. However, no mainstream politician took the lead in saying that the booming economic times might be an appropriate time for spending cuts or tax increases.

Discretionary Fiscal Policy: Summing Up

Expansionary fiscal policy can help to end recessions and contractionary fiscal policy can help to reduce inflation. Given the uncertainties over interest rate effects, time lags, temporary and permanent policies, and unpredictable political behavior, many economists and knowledgeable policymakers had concluded by the mid-1990s that discretionary fiscal policy was a blunt instrument, more like a club than a scalpel. It might still make sense to use it in extreme economic situations, like an especially deep or long recession. For less extreme situations, it was often preferable to let fiscal policy work through the automatic stabilizers and focus on monetary policy to steer short-term countercyclical efforts.

Key Concepts and Summary

Because fiscal policy affects the quantity of money that the government borrows in financial capital markets, it not only affects aggregate demand—it can also affect interest rates. If an expansionary fiscal policy also causes higher interest rates, then firms and households are discouraged from borrowing and spending, reducing aggregate demand in a situation called crowding out. Given the uncertainties over interest rate effects, time lags (implementation lag, legislative lag, and recognition lag), temporary and permanent policies, and unpredictable political behavior, many economists and knowledgeable policymakers have concluded that discretionary fiscal policy is a blunt instrument and better used only in extreme situations.

Self-Check Questions

Exercise:

Problem:

What would happen if expansionary fiscal policy was implemented in a recession but, due to lag, did not actually take effect until after the economy was back to potential GDP?

Solution:

Prices would be pushed up as a result of too much spending.

Exercise:

Problem:

What would happen if contractionary fiscal policy were implemented during an economic boom but, due to lag, it did not take effect until the economy slipped into recession?

Solution:

Employment would suffer as a result of too little spending.

Exercise:

Problem:

Do you think the typical time lag for fiscal policy is likely to be longer or shorter than the time lag for monetary policy? Explain your answer?

Solution:

Monetary policy probably has shorter time lags than fiscal policy. Imagine that the data becomes fairly clear that an economy is in or near a recession. Expansionary monetary policy can be carried out through open market operations, which can be done fairly quickly, since the Federal Reserve's Open Market Committee meets six times a year. Also, monetary policy takes effect through interest rates, which can change fairly quickly. However, fiscal policy is carried out through acts of Congress that need to be signed into law by the president. Negotiating such laws often takes months, and even after the laws are negotiated, it takes more months for spending programs or tax cuts to have an effect on the macroeconomy.

Review Questions**Exercise:****Problem:**

What are some practical weaknesses of discretionary fiscal policy?

Critical Thinking Questions**Exercise:****Problem:**

What is a potential problem with a temporary tax increase designed to increase aggregate demand if people know that it is temporary?

Exercise:

Problem:

If the government gives a \$300 tax cut to everyone in the country, explain the mechanism by which this will cause interest rates to rise.

References

Leduc, Sylvain, and Daniel Wilson. Federal Reserve Bank of San Francisco: Working Paper Series. "Are State Governments Roadblocks to Federal Stimulus? Evidence from Highway Grants in the 2009 Recovery Act. (Working Paper 2013-16)." Last modified July 2013. <http://www.frbsf.org/economic-research/files/wp2013-16.pdf>.

Lucking, Brian, and Daniel Wilson. "FRBSF Economic Letter-Fiscal Headwinds: Is the Other Shoe About to Drop?" *Federal Reserve Bank of San Francisco*. Last modified June 3, 2013. <http://www.frbsf.org/economic-research/publications/economic-letter/2013/june/fiscal-headwinds-federal-budget-policy/>.

Recovery.gov. "Track the Money."
<http://www.recovery.gov/Pages/default.aspx>.

Bastagli, Francesca, David Coady, and Sanjeev Gupta. International Monetary Fund. "IMF Staff Discussion Note: Income Inequality and Fiscal Policy." Last modified June 28, 2012. <http://www.imf.org/external/pubs/ft/sdn/2012/sdn1208.pdf>.

Glossary

crowding out

federal spending and borrowing causes interest rates to rise and business investment to fall

implementation lag

the time it takes for the funds relating to fiscal policy to be dispersed to the appropriate agencies to implement the programs

legislative lag

the time it takes to get a fiscal policy bill passed

recognition lag

the time it takes to determine that a recession has occurred

The Question of a Balanced Budget

By the end of this section, you will be able to:

- Understand the arguments for and against requiring the U.S. federal budget to be balanced
- Consider the long-run and short-run effects of a federal budget deficit

For many decades, going back to the 1930s, proposals have been put forward to require that the U.S. government balance its budget every year. In 1995, a proposed constitutional amendment that would require a balanced budget passed the U.S. House of Representatives by a wide margin, and failed in the U.S. Senate by only a single vote. (For the balanced budget to have become an amendment to the Constitution would have required a two-thirds vote by Congress and passage by three-quarters of the state legislatures.)

Most economists view the proposals for a perpetually balanced budget with bemusement. After all, in the short term, economists would expect the budget deficits and surpluses to fluctuate up and down with the economy and the automatic stabilizers. Economic recessions should automatically lead to larger budget deficits or smaller budget surpluses, while economic booms lead to smaller deficits or larger surpluses. A requirement that the budget be balanced each and every year would prevent these automatic stabilizers from working and would worsen the severity of economic fluctuations.

Some supporters of the balanced budget amendment like to argue that, since households must balance their own budgets, the government should too. But this analogy between household and government behavior is severely flawed. Most households do not balance their budgets every year. Some years households borrow to buy houses or cars or to pay for medical expenses or college tuition. Other years they repay loans and save funds in retirement accounts. After retirement, they withdraw and spend those savings. Also, the government is not a household for many reasons, one of which is that the government has macroeconomic responsibilities. The argument of Keynesian macroeconomic policy is that the government needs

to lean against the wind, spending when times are hard and saving when times are good, for the sake of the overall economy.

There is also no particular reason to expect a government budget to be balanced in the medium term of a few years. For example, a government may decide that by running large budget deficits, it can make crucial long-term investments in human capital and physical infrastructure that will build the long-term productivity of a country. These decisions may work out well or poorly, but they are not always irrational. Such policies of ongoing government budget deficits may persist for decades. As the U.S. experience from the end of World War II up to about 1980 shows, it is perfectly possible to run budget deficits almost every year for decades, but as long as the percentage increases in debt are smaller than the percentage growth of GDP, the debt/GDP ratio will decline at the same time.

Nothing in this argument should be taken as a claim that budget deficits are always a wise policy. In the short run, a government that runs a very large budget deficit can shift aggregate demand to the right and trigger severe inflation. Additionally, governments may borrow for foolish or impractical reasons. [The Macroeconomic Impacts of Government Borrowing](#) will discuss how large budget deficits, by reducing national saving, can in certain cases reduce economic growth and even contribute to international financial crises. A requirement that the budget be balanced in each calendar year, however, is a misguided overreaction to the fear that in some cases, budget deficits can become too large.

Note:**No Yellowstone Park?**

The federal budget shutdown of 2013 illustrated the many sides to fiscal policy and the federal budget. In 2013, Republicans and Democrats could not agree on which spending policies to fund and how large the government debt should be. Due to the severity of the recession in 2008–2009, the fiscal stimulus, and previous policies, the federal budget deficit and debt was historically high. One way to try to cut federal spending and borrowing was to refuse to raise the legal federal debt limit, or tie on conditions to appropriation bills to stop the Affordable Health Care Act.

This disagreement led to a two-week shutdown of the federal government and got close to the deadline where the federal government would default on its Treasury bonds. Finally, however, a compromise emerged and default was avoided. This shows clearly how closely fiscal policies are tied to politics.

Key Concepts and Summary

Balanced budget amendments are a popular political idea, but the economic merits behind such proposals are questionable. Most economists accept that fiscal policy needs to be flexible enough to accommodate unforeseen expenditures, such as wars or recessions. While persistent, large budget deficits can indeed be a problem, a balanced budget amendment prevents even small, temporary deficits that might, in some cases, be necessary.

Self-Check Questions

Exercise:

Problem:

How would a balanced budget amendment affect a decision by Congress to grant a tax cut during a recession?

Solution:

The government would have to make up the revenue either by raising taxes in a different area or cutting spending.

Exercise:

Problem:

How would a balanced budget amendment change the effect of automatic stabilizer programs?

Solution:

Programs where the amount of spending is not fixed, but rather determined by macroeconomic conditions, such as food stamps, would lose a great deal of flexibility if spending increases had to be met by corresponding tax increases or spending cuts.

Review Questions

Exercise:

Problem:

What are some of the arguments for and against a requirement that the federal government budget be balanced every year?

Critical Thinking Questions

Exercise:

Problem:

Do you agree or disagree with this statement: “It is in the best interest of our economy for Congress and the President to run a balanced budget each year.” Explain your answer.

Exercise:

Problem:

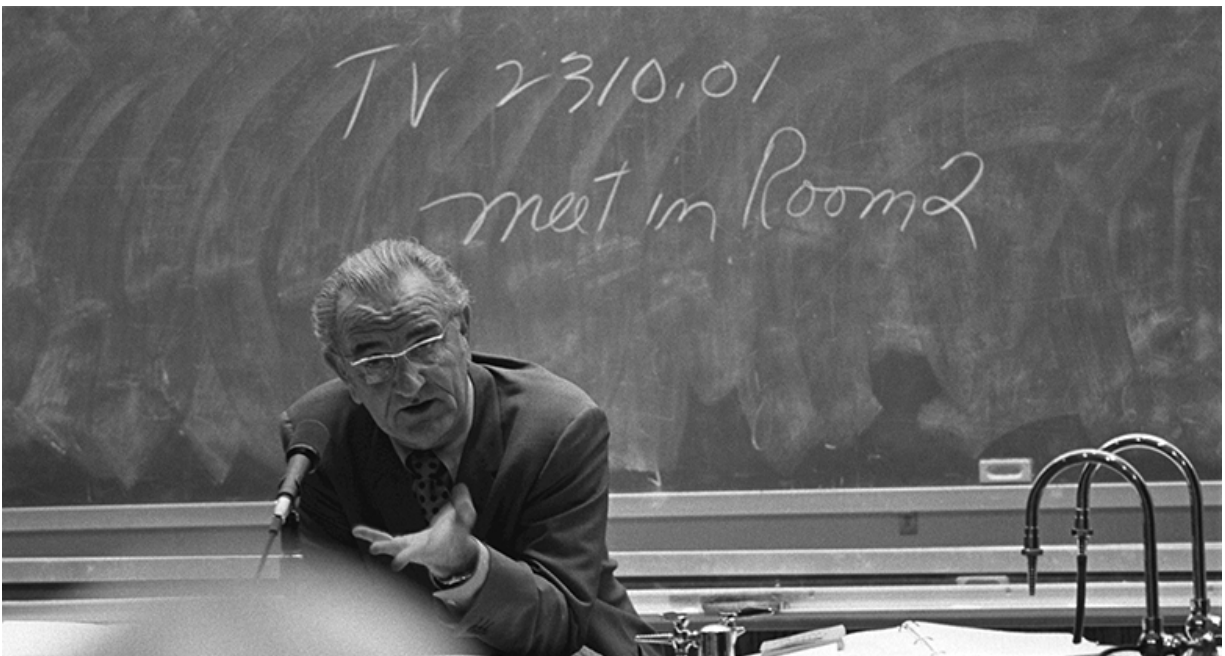
During the Great Recession of 2008–2009, what actions would have been required of Congress and the President had a balanced budget amendment to the Constitution been ratified? What impact would that have had on the unemployment rate?

Introduction to the Impacts of Government Borrowing

class="introduction"

President Lyndon B. Johnson

President
Lyndon
Johnson
played a
pivotal role
in financing
higher
education.
(Credit:
modificatio
n of image
by LBJ
Museum &
Library)



Note:**Financing Higher Education**

On November 8, 1965, President Lyndon B. Johnson signed The Higher Education Act of 1965 into law. With a stroke of the pen, he implemented what we know as the financial aid, work study, and student loan programs to help Americans pay for a college education. In his remarks, the President said:

"Here the seeds were planted from which grew my firm conviction that for the individual, education is the path to achievement and fulfillment; for the Nation, it is a path to a society that is not only free but civilized; and for the world, it is the path to peace—for it is education that places reason over force."

This Act, he said, "is responsible for funding higher education for millions of Americans. It is the embodiment of the United States' investment in 'human capital'." Since the Act was first signed into law, it has been renewed several times.

The purpose of The Higher Education Act of 1965 was to build the country's human capital by creating educational opportunity for millions of Americans. The three criteria used to judge eligibility are income, full-time or part-time attendance, and the cost of the institution. According to the 2011–2012 National Postsecondary Student Aid Study (NPSAS:12), in the 2011–2012 school year, over 70% of all full-time college students received some form of federal financial aid; 47% received grants; and another 55% received federal government student loans. The budget to support financial aid has increased not only because of increased enrollment, but also because of increased tuition and fees for higher education. These increases are currently being questioned. The President and Congress are charged with balancing fiscal responsibility and important government-financed expenditures like investing in human capital.

Note:**Introduction to the Impacts of Government Borrowing**

In this chapter, you will learn about:

- How Government Borrowing Affects Investment and the Trade Balance
- Fiscal Policy, Investment, and Economic Growth
- How Government Borrowing Affects Private Saving
- Fiscal Policy and the Trade Balance

Governments have many competing demands for financial support. Any spending should be tempered by fiscal responsibility and by looking carefully at the spending's impact. When a government spends more than it collects in taxes, it runs a budget deficit. It then needs to borrow. When government borrowing becomes especially large and sustained, it can substantially reduce the financial capital available to private sector firms, as well as lead to trade imbalances and even financial crises.

The [Government Budgets and Fiscal Policy](#) chapter introduced the concepts of deficits and debt, as well as how a government could use fiscal policy to address recession or inflation. This chapter begins by building on the national savings and investment identity, first introduced in [The International Trade and Capital Flows](#) chapter, to show how government borrowing affects firms' physical capital investment levels and trade balances. A prolonged period of budget deficits may lead to lower economic growth, in part because the funds borrowed by the government to fund its budget deficits are typically no longer available for private investment. Moreover, a sustained pattern of large budget deficits can lead to disruptive economic patterns of high inflation, substantial inflows of financial capital from abroad, plummeting exchange rates, and heavy strains on a country's banking and financial system.

How Government Borrowing Affects Investment and the Trade Balance

By the end of this section, you will be able to:

- Explain the national saving and investment identity in terms of demand and supply
- Evaluate the role of budget surpluses and trade surpluses in national saving and investment identity

When governments are borrowers in financial markets, there are three possible sources for the funds from a macroeconomic point of view: (1) households might save more; (2) private firms might borrow less; and (3) the additional funds for government borrowing might come from outside the country, from foreign financial investors. Let's begin with a review of why one of these three options must occur, and then explore how interest rates and exchange rates adjust to these connections.

The National Saving and Investment Identity

The national saving and investment identity, first introduced in [The International Trade and Capital Flows](#) chapter, provides a framework for showing the relationships between the sources of demand and supply in financial capital markets. The identity begins with a statement that must always hold true: the quantity of financial capital supplied in the market must equal the quantity of financial capital demanded.

The U.S. economy has two main sources for financial capital: private savings from inside the U.S. economy and public savings.

Equation:

$$\text{Total savings} = \text{Private savings (S)} + \text{Public savings (T - G)}$$

These include the inflow of foreign financial capital from abroad. The inflow of savings from abroad is, by definition, equal to the trade deficit, as explained in [The International Trade and Capital Flows](#) chapter. So this inflow of foreign investment capital can be written as imports (M) minus exports (X). There are also two main sources of demand for financial capital: private sector investment (I) and government borrowing. Government borrowing in any given year is equal to the budget deficit, and can be written as the difference between government spending (G) and net taxes (T). Let's call this equation 1.

Equation:

$$\begin{aligned} \text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ \text{Private savings} + \text{Inflow of foreign savings} &= \text{Private investment} + \text{Government budget deficit} \\ S + (M - X) &= I + (G - T) \end{aligned}$$

Governments often spend more than they receive in taxes and, therefore, public savings (T - G) is negative. This causes a need to borrow money in the amount of (G - T) instead of adding to the nation's savings. If this is the case, governments can be viewed as demanders of financial capital instead of suppliers. So, in algebraic terms, the national savings and investment identity can be rewritten like this:

Equation:

$$\begin{aligned} \text{Private investment} &= \text{Private savings} + \text{Public savings} + \text{Trade deficit} \\ I &= S + (T - G) + (M - X) \end{aligned}$$

Let's call this equation 2. A change in any part of the national saving and investment identity must be accompanied by offsetting changes in at least one other part of the equation because the equality of quantity supplied and quantity demanded is always assumed to hold. If the government budget deficit changes, then either private saving or investment or the trade balance—or some combination of the three—must change as well. [\[link\]](#) shows the possible effects.

Effects of Change in Budget Surplus or Deficit on Investment, Savings, and The Trade Balance

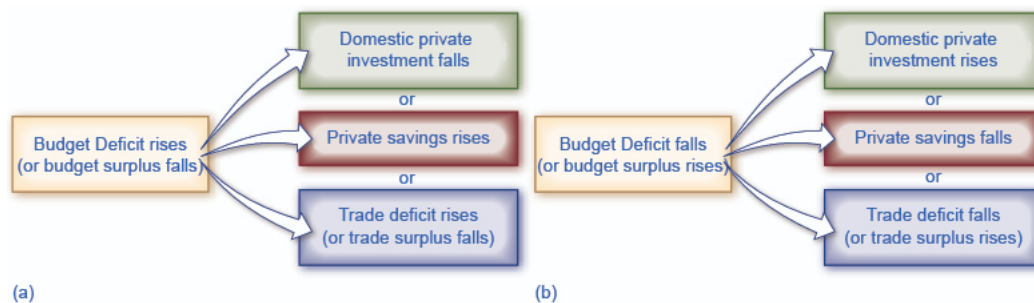


Chart (a) shows the potential results when the budget deficit rises (or budget surplus falls).
 Chart (b) shows the potential results when the budget deficit falls (or budget surplus rises).

What about Budget Surpluses and Trade Surpluses?

The national saving and investment identity must always hold true because, by definition, the quantity supplied and quantity demanded in the financial capital market must always be equal. However, the formula will look somewhat different if the government budget is in deficit rather than surplus or if the balance of trade is in surplus rather than deficit. For example, in 1999 and 2000, the U.S. government had budget surpluses, although the economy was still experiencing trade deficits. When the government was running budget surpluses, it was acting as a saver rather than a borrower, and supplying rather than demanding financial capital. As a result, the national saving and investment identity during this time would be more properly written:

Equation:

$$\begin{aligned} \text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ \text{Private savings} + \text{Trade deficit} + \text{Government surplus} &= \text{Private investment} \\ S + (M - X) + (T - G) &= I \end{aligned}$$

Let's call this equation 3. Notice that this expression is mathematically the same as equation 2 except the savings and investment sides of the identity have simply flipped sides.

During the 1960s, the U.S. government was often running a budget deficit, but the economy was typically running trade surpluses. Since a trade surplus means that an economy is experiencing a net outflow of financial capital, the national saving and investment identity would be written:

Equation:

$$\begin{aligned} \text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ \text{Private savings} &= \text{Private investment} + \text{Outflow of foreign savings} + \text{Government} \\ S &= I + (X - M) + (G - T) \end{aligned}$$

Instead of the balance of trade representing part of the supply of financial capital, which occurs with a trade deficit, a trade surplus represents an outflow of financial capital leaving the domestic economy and being invested elsewhere in the world.

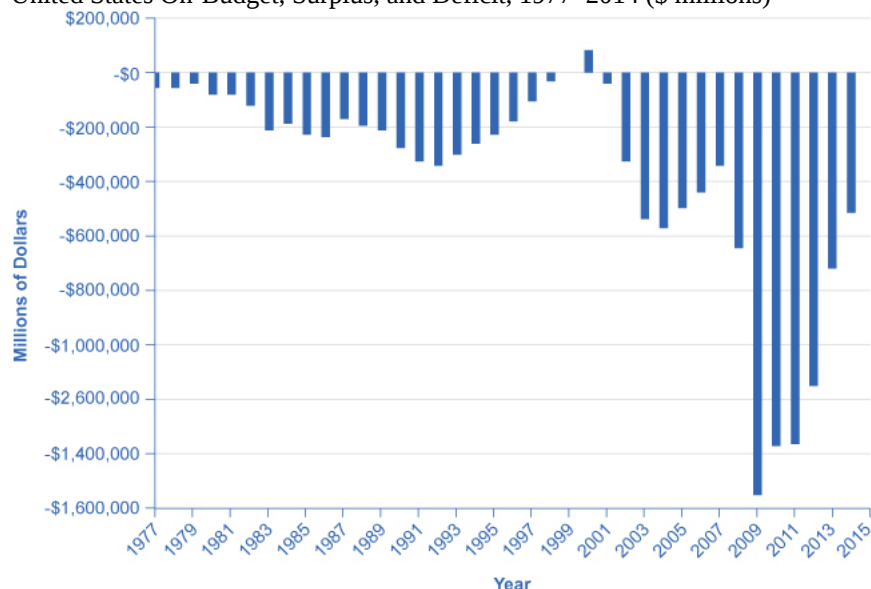
Equation:

$$\begin{aligned} \text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ \text{Private savings} &= \text{Private investment} + \text{Government budget deficit} + \text{Trade surplus} \\ S &= I + (G - T) + (X - M) \end{aligned}$$

The point to this parade of equations is that the national saving and investment identity is assumed to always hold. So when you write these relationships, it is important to engage your brain and think about what is on the supply side and what is on the demand side of the financial capital market before you put pencil to paper.

As can be seen in [\[link\]](#), the Office of Management and Budget shows that the United States has consistently run budget deficits since 1977, with the exception of 1999 and 2000. What is alarming is the dramatic increase in budget deficits that has occurred since 2008, which in part reflects declining tax revenues and increased safety net expenditures due to the Great Recession. (Recall that T is net taxes. When the government must transfer funds back to individuals for safety net expenditures like Social Security and unemployment benefits, budget deficits rise.) These deficits have implications for the future health of the U.S. economy.

United States On-Budget, Surplus, and Deficit, 1977–2014 (\$ millions)



The United States has run a budget deficit for over 30 years, with the exception of 1999 and 2000. Military expenditures, entitlement programs, and the decrease in tax revenue coupled with increased safety net support during the Great Recession are major contributors to the dramatic increases in the deficit after 2008. (Source: Table 1.1, "Summary of Receipts, Outlays, and Surpluses or Deficits," <https://www.whitehouse.gov/omb/budget/Historicals>)

A rising budget deficit may result in a fall in domestic investment, a rise in private savings, or a rise in the trade deficit. The following modules discuss each of these possible effects in more detail.

Key Concepts and Summary

A change in any part of the national saving and investment identity suggests that if the government budget deficit changes, then either private savings, private investment in physical capital, or the trade balance—or some combination of the three—must change as well.

Self-Check Questions

Exercise:

Problem:

In a country, private savings equals 600, the government budget surplus equals 200, and the trade surplus equals 100. What is the level of private investment in this economy?

Solution:

We use the national savings and investment identity to solve this question. In this case, the government has a budget surplus, so the government surplus appears as part of the supply of financial capital. Then:

Equation:

$$\begin{aligned}\text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ S + (T - G) &= I + (X - M) \\ 600 + 200 &= I + 100 \\ I &= 700\end{aligned}$$

Exercise:**Problem:**

Assume an economy has a budget surplus of 1,000, private savings of 4,000, and investment of 5,000.

- Write out a national saving and investment identity for this economy.
 - What will be the balance of trade in this economy?
 - If the budget surplus changes to a budget deficit of 1000, with private saving and investment unchanged, what is the new balance of trade in this economy?
-

Solution:

- Since the government has a budget surplus, the government budget term appears with the supply of capital. The following shows the national savings and investment identity for this economy.

Equation:

$$\begin{aligned}\text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ S + (T - G) &= I + (X - M)\end{aligned}$$

- Plugging the given values into the identity shown in part (a), we find that $(X - M) = 0$.
- Since the government has a budget deficit, the government budget term appears with the demand for capital. You do not know in advance whether the economy has a trade deficit or a trade surplus. But when you see that the quantity demanded of financial capital exceeds the quantity supplied, you know that there must be an additional quantity of financial capital supplied by foreign investors, which means a trade deficit of 2000. This example shows that in this case there is a higher budget deficit, and a higher trade deficit.

Equation:

$$\begin{aligned}\text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ S + (M - X) &= I + (G - T) \\ 4000 + 2000 &= 5000 + 1000\end{aligned}$$

Review Questions**Exercise:**

Problem:

Based on the national saving and investment identity, what are the three ways the macroeconomy might react to greater government budget deficits?

Exercise:

Problem:

How would you expect larger budget deficits to affect private sector investment in physical capital? Why?

Critical Thinking Questions

Exercise:

Problem:

Assume there is no discretionary increase in government spending. Explain how an improving economy will affect the budget balance and, in turn, investment and the trade balance.

Exercise:

Problem:

Explain how decreased domestic investments that occur due to a budget deficit will affect future economic growth.

Exercise:

Problem:

The U.S. government has shut down a number of times in recent history. Explain how a government shutdown will affect the variables in the national investment and savings identity. Could the shutdown affect the government budget deficit?

Fiscal Policy, Investment, and Economic Growth

By the end of this section, you will be able to:

- Explain crowding out and its effect on physical capital investment
- Explain the relationship between budget deficits and interest rates
- Identify why economic growth is tied to investments in physical capital, human capital, and technology

The underpinnings of economic growth are investments in physical capital, human capital, and technology, all set in an economic environment where firms and individuals can react to the incentives provided by well-functioning markets and flexible prices. Government borrowing can reduce the financial capital available for private firms to invest in physical capital. But government spending can also encourage certain elements of long-term growth, such as spending on roads or water systems, on education, or on research and development that creates new technology.

Crowding Out Physical Capital Investment

A larger budget deficit will increase demand for financial capital. If private saving and the trade balance remain the same, then less financial capital will be available for private investment in physical capital. When government borrowing soaks up available financial capital and leaves less for private investment in physical capital, the result is known as crowding out.

To understand the potential impact of crowding out, consider the situation of the U.S. economy before the exceptional circumstances of the recession that started in late 2007. In 2005, for example, the budget deficit was roughly 4% of GDP. Private investment by firms in the U.S. economy has hovered in the range of 14% to 18% of GDP in recent decades. However, in any given year, roughly half of U.S. investment in physical capital just replaces machinery and equipment that has worn out or become technologically obsolete. Only about half represents an increase in the total quantity of physical capital in the economy. So investment in new physical capital in any year is about 7% to 9% of GDP. In this situation, even U.S.

budget deficits in the range of 4% of GDP can potentially crowd out a substantial share of new investment spending. Conversely, a smaller budget deficit (or an increased budget surplus) increases the pool of financial capital available for private investment.

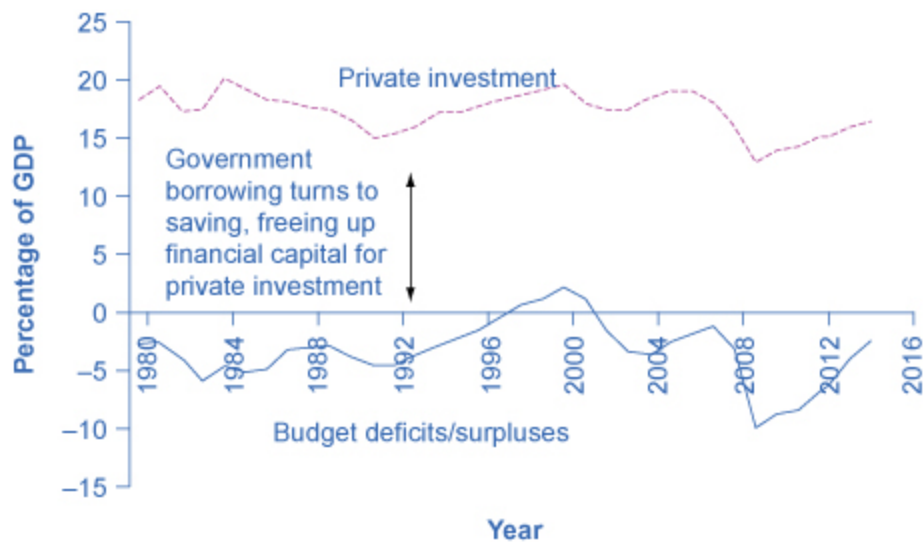
Note:

Visit this [website](#) to view the “U.S. Debt Clock.”



The patterns of U.S. budget deficits and private investment since 1980 are shown in [\[link\]](#). If greater government deficits lead to less private investment in physical capital, and reduced government deficits or budget surpluses lead to more investment in physical capital, these two lines should move up and down at the same time. This pattern occurred in the late 1990s and early 2000s. The U.S. federal budget went from a deficit of 2.2% of GDP in 1995 to a budget surplus of 2.4% of GDP in 2000—a swing of 4.6% of GDP. From 1995 to 2000, private investment in physical capital rose from 15% to 18% of GDP—a rise of 3% of GDP. Then, when the U.S. government again started running budget deficits in the early 2000s, less financial capital became available for private investment, and the rate of private investment fell back to about 15% of GDP by 2003.

U.S. Budget Deficits/Surpluses and Private Investment



The connection between private savings and flows of international capital plays a role in budget deficits and surpluses. Consequently, government borrowing and private investment sometimes rise and fall together. For example, the 1990s show a pattern in which reduced government borrowing helped to reduce crowding out so that more funds were available for private investment.

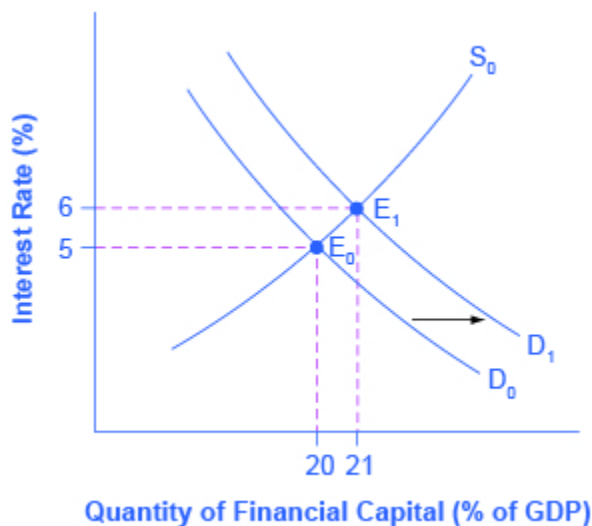
This argument does not claim that a government's budget deficits will exactly shadow its national rate of private investment; after all, private saving and inflows of foreign financial investment must also be taken into account. In the mid-1980s, for example, government budget deficits increased substantially without a corresponding drop off in private investment. In 2009, nonresidential private fixed investment dropped by \$300 billion from its previous level of \$1,941 billion in 2008, primarily because, during a recession, firms lack both the funds and the incentive to invest. Investment growth between 2009 and 2014 averaged approximately 5.9% to \$2,210.5 billion—only slightly above its 2008 level, according to the Bureau of Economic Analysis. During that same period, interest rates dropped from 3.94% to less than a quarter percent as the Federal Reserve took dramatic action to prevent a depression by increasing the money

supply through lowering short-term interest rates. The "crowding out" of private investment due to government borrowing to finance expenditures appears to have been suspended during the Great Recession. However, as the economy improves and interest rates rise, borrowing by the government may potentially create pressure on interest rates.

The Interest Rate Connection

Assume that government borrowing of substantial amounts will have an effect on the quantity of private investment. How will this affect interest rates in financial markets? In [\[link\]](#), the original equilibrium (E_0) where the demand curve (D_0) for financial capital intersects with the supply curve (S_0) occurs at an interest rate of 5% and an equilibrium quantity equal to 20% of GDP. However, as the government budget deficit increases, the demand curve for financial capital shifts from D_0 to D_1 . The new equilibrium (E_1) occurs at an interest rate of 6% and an equilibrium quantity of 21% of GDP.

Budget Deficits and Interest Rates



In the financial market, an increase in government borrowing can shift the demand curve for financial capital to the right from D_0 to D_1 . As the equilibrium interest rate shifts from E_0 to E_1 , the interest rate rises from 5% to 6% in this example. The higher interest rate is one

economic mechanism by which
government borrowing can crowd out
private investment.

A survey of economic studies on the connection between government borrowing and interest rates in the U.S. economy suggests that an increase of 1% in the budget deficit will lead to a rise in interest rates of between 0.5 and 1.0%, other factors held equal. In turn, a higher interest rate tends to discourage firms from making physical capital investments. One reason government budget deficits crowd out private investment, therefore, is the increase in interest rates. There are, however, economic studies that show a limited connection between the two (at least in the United States), but as the budget deficit grows, the dangers of rising interest rates become more real.

At this point, you may wonder about the Federal Reserve. After all, can the Federal Reserve not use expansionary monetary policy to reduce interest rates, or in this case, to prevent interest rates from rising? This useful question emphasizes the importance of considering how fiscal and monetary policies work in relation to each other. Imagine a central bank faced with a government that is running large budget deficits, causing a rise in interest rates and crowding out private investment. If the budget deficits are increasing aggregate demand when the economy is already producing near potential GDP, threatening an inflationary increase in price levels, the central bank may react with a contractionary monetary policy. In this situation, the higher interest rates from the government borrowing would be made even higher by contractionary monetary policy, and the government borrowing might crowd out a great deal of private investment.

On the other hand, if the budget deficits are increasing aggregate demand when the economy is producing substantially less than potential GDP, an inflationary increase in the price level is not much of a danger and the central bank might react with expansionary monetary policy. In this situation, higher interest rates from government borrowing would be largely offset by lower interest rates from expansionary monetary policy, and there would be little crowding out of private investment.

However, even a central bank cannot erase the overall message of the national savings and investment identity. If government borrowing rises, then private investment must fall, or private saving must rise, or the trade deficit must fall. By reacting with contractionary or expansionary monetary policy, the central bank can only help to determine which of these outcomes is likely.

Public Investment in Physical Capital

Government can invest in physical capital directly: roads and bridges; water supply and sewers; seaports and airports; schools and hospitals; plants that generate electricity, like hydroelectric dams or windmills; telecommunications facilities; and weapons used by the military. In 2014, the U.S. federal government budget for Fiscal Year 2014 shows that the United States spent about \$92 billion on transportation, including highways, mass transit, and airports. [\[link\]](#) shows the total outlay for 2014 for major public physical capital investment by the federal government in the United States. Physical capital related to the military or to residences where people live is omitted from this table, because the focus here is on public investments that have a direct effect on raising output in the private sector.

Type of Public Physical Capital	Federal Outlays 2014 (\$ millions)
Transportation	\$91,915
Community and regional development	\$20,670
Natural resources and the environment	\$36,171

Type of Public Physical Capital	Federal Outlays 2014 (\$ millions)
Education, training, employment, and social services	\$90,615
Other	\$37,282
Total	\$276,653

Grants for Major Physical Capital Investment, 2014

Public physical capital investment of this sort can increase the output and productivity of the economy. An economy with reliable roads and electricity will be able to produce more. But it is hard to quantify how much government investment in physical capital will benefit the economy, because government responds to political as well as economic incentives. When a firm makes an investment in physical capital, it is subject to the discipline of the market: If it does not receive a positive return on investment, the firm may lose money or even go out of business.

In some cases, lawmakers make investments in physical capital as a way of spending money in the districts of key politicians. The result may be unnecessary roads or office buildings. Even if a project is useful and necessary, it might be done in a way that is excessively costly, because local contractors who make campaign contributions to politicians appreciate the extra business. On the other hand, governments sometimes do not make the investments they should because a decision to spend on infrastructure does not need to just make economic sense; it must be politically popular as well. Managing public investment so that it is done in a cost-effective way can be difficult.

If a government decides to finance an investment in public physical capital with higher taxes or lower government spending in other areas, it need not worry that it is directly crowding out private investment. Indirectly however, higher household taxes could cut down on the level of private savings available and have a similar effect. If a government decides to

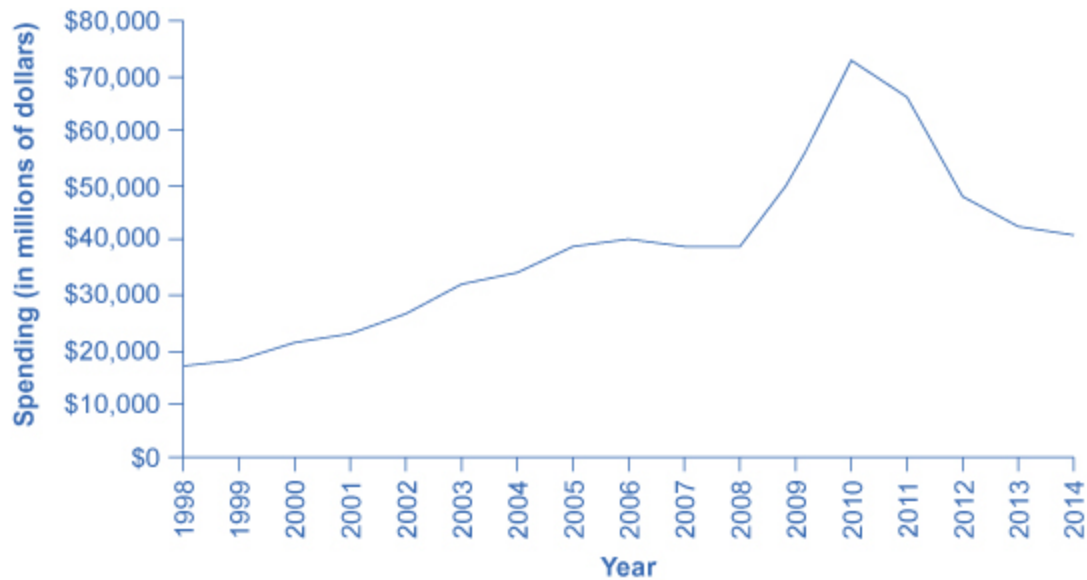
finance an investment in public physical capital by borrowing, it may end up increasing the quantity of public physical capital at the cost of crowding out investment in private physical capital, which is more beneficial to the economy would be dependent on the project being considered.

Public Investment in Human Capital

In most countries, the government plays a large role in society's investment in human capital through the education system. A highly educated and skilled workforce contributes to a higher rate of economic growth. For the low-income nations of the world, additional investment in human capital seems likely to increase productivity and growth. For the United States, tough questions have been raised about how much increases in government spending on education will improve the actual level of education.

Among economists, discussions of education reform often begin with some uncomfortable facts. As shown in [\[link\]](#), spending per student for kindergarten through grade 12 (K–12) increased substantially in real dollars through 2010. The U.S. Census Bureau reports that current spending per pupil for elementary and secondary education rose from \$5,001 in 1998 to \$10,608 in 2012. However, as measured by standardized tests like the SAT, the level of student academic achievement has barely budged in recent decades. Indeed, on international tests, U.S. students lag behind students from many other countries. (Of course, test scores are an imperfect measure of education for a variety of reasons. It would be difficult, however, to argue that there are not real problems in the U.S. education system and that the tests are just inaccurate.)

Total Spending for Elementary, Secondary, and Vocational Education (1998–2014) in the United States



The graph shows that government spending on education was continually increasing up until 2006 where it leveled off until 2008 when it increased dramatically. Since 2010, spending has steadily decreased. (Source: Office of Management and Budget)

The fact that increased financial resources have not brought greater measurable gains in student performance has led some education experts to question whether the problems may be due to structure, not just to the resources spent.

Other government programs seek to increase human capital either before or after the K–12 education system. Programs for early childhood education, like the federal **Head Start program**, are directed at families where the parents may have limited educational and financial resources. Government also offers substantial support for universities and colleges. For example, in the United States about 60% of students take at least a few college or university classes beyond the high school level. In Germany and Japan, about half of all students take classes beyond the comparable high school level. In the countries of Latin America, only about one student in four takes classes beyond the high school level, and in the nations of sub-Saharan Africa, only about one student in 20.

Not all spending on educational human capital needs to happen through the government: many college students in the United States pay a substantial share of the cost of their education. If low-income countries of the world are going to experience a widespread increase in their education levels for grade-school children, government spending seems likely to play a substantial role. For the U.S. economy, and for other high-income countries, the primary focus at this time is more on how to get a bigger return from existing spending on education and how to improve the performance of the average high school graduate, rather than dramatic increases in education spending.

How Fiscal Policy Can Improve Technology

Research and development (R&D) efforts are the lifeblood of new technology. According to the National Science Foundation, federal outlays for research, development, and physical plant improvements to various governmental agencies have remained at an average of 8.8% of GDP. About one-fifth of U.S. R&D spending goes to defense and space-oriented research. Although defense-oriented R&D spending may sometimes produce consumer-oriented spinoffs, R&D that is aimed at producing new weapons is less likely to benefit the civilian economy than direct civilian R&D spending.

Fiscal policy can encourage R&D using either direct spending or tax policy. Government could spend more on the R&D that is carried out in government laboratories, as well as expanding federal R&D grants to universities and colleges, nonprofit organizations, and the private sector. By 2014, the federal share of R&D outlays totaled \$135.5 billion, or about 4% of the federal government's total budget outlays, according to data from the National Science Foundation. Fiscal policy can also support R&D through tax incentives, which allow firms to reduce their tax bill as they increase spending on research and development.

Summary of Fiscal Policy, Investment, and Economic Growth

Investment in physical capital, human capital, and new technology is essential for long-term economic growth, as summarized in [\[link\]](#). In a

market-oriented economy, private firms will undertake most of the investment in physical capital, and fiscal policy should seek to avoid a long series of outsized budget deficits that might crowd out such investment. The effects of many growth-oriented policies will be seen very gradually over time, as students are better educated, physical capital investments are made, and new technologies are invented and implemented.

	Physical Capital	Human Capital	New Technology
Private Sector	New investment in property and equipment	On-the-job training	Research and development
Public Sector	Public infrastructure	Public education Job training	Research and development encouraged through private sector incentives and direct spending.

Investment Role of Public and Private Sector in a Market Economy

Key Concepts and Summary

Economic growth comes from a combination of investment in physical capital, human capital, and technology. Government borrowing can crowd out private sector investment in physical capital, but fiscal policy can also increase investment in publicly owned physical capital, human capital (education), and research and development. Possible methods for improving

education and society's investment in human capital include spending more money on teachers and other educational resources, and reorganizing the education system to provide greater incentives for success. Methods for increasing research and development spending to generate new technology include direct government spending on R&D and tax incentives for businesses to conduct additional R&D.

Self-Check Questions

Exercise:

Problem:

Why have many education experts recently placed an emphasis on altering the incentives faced by U.S. schools rather than on increasing their budgets? Without endorsing any of these proposals as especially good or bad, list some of the ways in which incentives for schools might be altered.

Solution:

In the last few decades, spending per student has climbed substantially. However, test scores have fallen over this time. This experience has led a number of experts to argue that the problem is not resources—or is not just resources by itself—but is also a problem of how schools are organized and managed and what incentives they have for success. There are a number of proposals to alter the incentives that schools face, but relatively little hard evidence on what proposals work well. Without trying to evaluate whether these proposals are good or bad ideas, you can just list some of them: testing students regularly; rewarding teachers or schools that perform well on such tests; requiring additional teacher training; allowing students to choose between public schools; allowing teachers and parents to start new schools; giving student “vouchers” that they can use to pay tuition at either public or private schools.

Exercise:

Problem:

What are some steps the government can take to encourage research and development?

Solution:

The government can direct government spending to R&D. It can also create tax incentives for business to invest in R&D.

Review Questions**Exercise:****Problem:**

What are some of the ways fiscal policy might encourage economic growth?

Exercise:**Problem:**

What are some fiscal policies for improving a society's human capital?

Exercise:**Problem:**

What are some fiscal policies for improving the technologies that the economy will have to draw upon in the future?

Exercise:**Problem:**

Explain how cuts in funding for programs such as Head Start might affect the development of human capital in the United States.

Critical Thinking Questions

Exercise:**Problem:**

Explain why the government might prefer to provide incentives to private firms to do investment or research and development, rather than simply doing the spending itself?

Exercise:**Problem:**

Under what condition would crowding out not inhibit long-run economic growth? Under what condition would crowding out impede long-run economic growth?

Exercise:**Problem:**

What must take place for the government to run deficits without any crowding out?

Problems**Exercise:****Problem:**

During the most recent recession, some economists argued that the change in the interest rates that comes about due to deficit spending implied in the demand and supply of financial capital graph would not occur. A simple reason was that the government was stepping in to invest when private firms were not. Using a graph, explain how the deficit demand is offset by the use by government in investment.

References

U.S. Department of Commerce: Bureau of Economic Analysis. “National Data: National Income and Product Accounts Tables.” Accessed December 1, 2013. <http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&910=X&911=0&903=146&904=2008&905=2013&906=A>.

Board of Governors of the Federal Reserve System, “Selected Interest Rates (Daily) – H.15.” Accessed December 10, 2013. <http://www.federalreserve.gov/releases/h15/data.htm>.

The White House. “Fiscal Year 2013 Historical Tables: Budget of the U.S. Government.” Accessed December 12, 2013. <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/hist.pdf>.

The National Science Foundation. Accessed December 19, 2013. <http://www.nsf.gov/>.

Glossary

Head Start program

a program for early childhood education directed at families with limited educational and financial resources.

How Government Borrowing Affects Private Saving

By the end of this section, you will be able to:

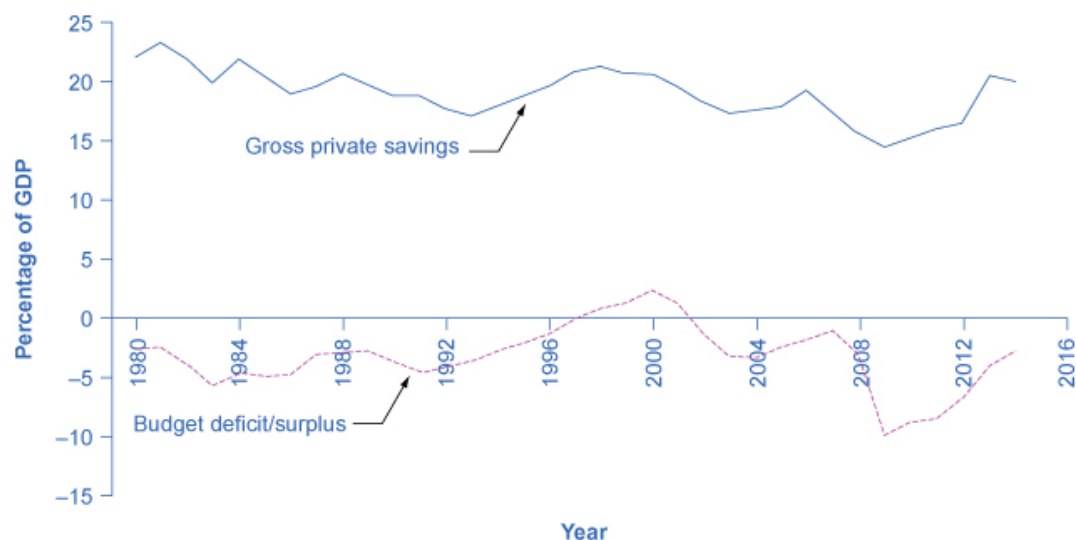
- Apply Ricardian equivalence to evaluate how government borrowing affects private saving
- Interpret a graphic representation of Ricardian equivalence

A change in government budgets may impact private saving. Imagine that people watch government budgets and adjust their savings accordingly. For example, whenever the government runs a budget deficit, people might reason: “Well, a higher budget deficit means that I’m just going to owe more taxes in the future to pay off all that government borrowing, so I’ll start saving now.” If the government runs budget surpluses, people might reason: “With these budget surpluses (or lower budget deficits), interest rates are falling, so that saving is less attractive. Moreover, with a budget surplus the country will be able to afford a tax cut sometime in the future. I won’t bother saving as much now.”

The theory that rational private households might shift their saving to offset government saving or borrowing is known as **Ricardian equivalence** because the idea has intellectual roots in the writings of the early nineteenth-century economist David Ricardo (1772–1823). If Ricardian equivalence holds completely true, then in the national saving and investment identity, any change in budget deficits or budget surpluses would be completely offset by a corresponding change in private saving. As a result, changes in government borrowing would have no effect at all on either physical capital investment or trade balances.

In practice, the private sector only sometimes and partially adjusts its savings behavior to offset government budget deficits and surpluses. [\[link\]](#) shows the patterns of U.S. government budget deficits and surpluses and the rate of private saving—which includes saving by both households and firms—since 1980. The connection between the two is not at all obvious. In the mid-1980s, for example, government budget deficits were quite large, but there is no corresponding surge of private saving. However, when budget deficits turn to surpluses in the late 1990s, there is a simultaneous decline in private saving. When budget deficits get very large in 2008 and 2009, on the other hand, there is some sign of a rise in saving. A variety of statistical studies based on the U.S. experience suggests that when government borrowing increases by \$1, private saving rises by about 30 cents. A World Bank study done in the late 1990s, looking at government budgets and private saving behavior in countries around the world, found a similar result.

U.S. Budget Deficits and Private Savings



The theory of Ricardian equivalence suggests that any increase in government borrowing will be offset by additional private saving, while any decrease in government borrowing will be offset by reduced private saving. Sometimes this theory holds true, and sometimes it does not hold true at all. (Source: Bureau of Economic Analysis and Federal Reserve Economic Data)

So private saving does increase to some extent when governments run large budget deficits, and private saving falls when governments reduce deficits or run large budget surpluses. However, the offsetting effects of private saving compared to government borrowing are much less than one-to-one. In addition, this effect can vary a great deal from country to country, from time to time, and over the short run and the long run.

If the funding for a larger budget deficit comes from international financial investors, then a budget deficit may be accompanied by a trade deficit. In some countries, this pattern of a **twin deficits** has set the stage for international financial investors first to send their funds to a country and cause an appreciation of its exchange rate and then to pull their funds out and cause a depreciation of the exchange rate and a financial crisis as well. It depends on whether funding comes from international financial investors.

Key Concepts and Summary

The theory of Ricardian equivalence holds that changes in government borrowing or saving will be offset by changes in private saving. Thus, higher budget deficits will be offset by greater private saving, while larger budget surpluses will be offset by greater private borrowing. If the theory holds true, then changes in government borrowing or saving

would have no effect on private investment in physical capital or on the trade balance. However, empirical evidence suggests that the theory holds true only partially.

Self-Check Questions

Exercise:

Problem:

Imagine an economy in which Ricardian equivalence holds. This economy has a budget deficit of 50, a trade deficit of 20, private savings of 130, and investment of 100. If the budget deficit rises to 70, how are the other terms in the national saving and investment identity affected?

Solution:

Ricardian equivalence means that private saving changes to offset exactly any changes in the government budget. So, if the deficit increases by 20, private saving increases by 20 as well, and the trade deficit and the budget deficit will not change from their original levels. The original national saving and investment identity is written below. Notice that if any change in the $(G - T)$ term is offset by a change in the S term, then the other terms do not change. So if $(G - T)$ rises by 20, then S must also increase by 20.

Equation:

$$\begin{aligned}\text{Quantity supplied of financial capital} &= \text{Quantity demanded of financial capital} \\ S + (M - X) &= I + (G - T) \\ 130 + 20 &= 100 + 50\end{aligned}$$

Review Questions

Exercise:

Problem: What is the theory of Ricardian equivalence?

Exercise:

Problem:

What does the concept of rationality have to do with Ricardian equivalence?

Critical Thinking Questions

Exercise:**Problem:**

Explain whether or not you agree with the premise of the Ricardian equivalence theory that rational people might reason: “Well, a higher budget deficit (surplus) means that I’m just going to owe more (less) taxes in the future to pay off all that government borrowing, so I’ll start saving (spending) now.” Why or why not?

Problems**Exercise:****Problem:**

Illustrate the concept of Ricardian equivalence using the demand and supply of financial capital graph.

Glossary

Ricardian equivalence

the theory that rational private households might shift their saving to offset government saving or borrowing

twin deficits

deficits that occur when a country is running both a trade and a budget deficit

Fiscal Policy and the Trade Balance

By the end of this section, you will be able to:

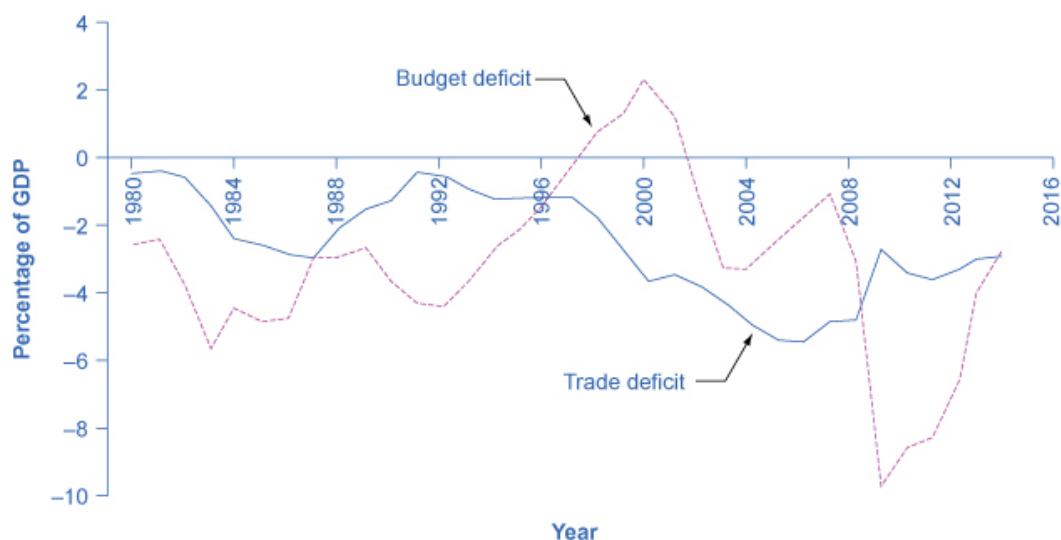
- Discuss twin deficits as they related to budget and trade deficit
- Explain the relationship between budget deficits and exchange rates
- Explain the relationship between budget deficits and inflation
- Identify causes of recessions

Government budget balances can affect the trade balance. As [The Keynesian Perspective](#) chapter discusses, a net inflow of foreign financial investment always accompanies a trade deficit, while a net outflow of financial investment always accompanies a trade surplus. One way to understand the connection from budget deficits to trade deficits is that when government creates a budget deficit with some combination of tax cuts or spending increases, it will increase aggregate demand in the economy, and some of that increase in aggregate demand will result in a higher level of imports. A higher level of imports, with exports remaining fixed, will cause a larger trade deficit. That means foreigners' holdings of dollars increase as Americans purchase more imported goods. Foreigners use those dollars to invest in the United States, which leads to an inflow of foreign investment. One possible source of funding our budget deficit is foreigners buying Treasury securities that are sold by the U.S. government. So a budget deficit is often accompanied by a trade deficit.

Twin Deficits?

In the mid-1980s, it was common to hear economists and even newspaper articles refer to the twin deficits, as the budget deficit and trade deficit both grew substantially. [\[link\]](#) shows the pattern. The federal budget deficit went from 2.6% of GDP in 1981 to 5.1% of GDP in 1985—a drop of 2.5% of GDP. Over that time, the trade deficit moved from 0.5% in 1981 to 2.9% in 1985—a drop of 2.4% of GDP. In the mid-1980s, the considerable increase in government borrowing was matched by an inflow of foreign investment capital, so the government budget deficit and the trade deficit moved together.

U.S. Budget Deficits and Trade Deficits



In the 1980s, the budget deficit and the trade deficit declined at the same time. However, since then, the deficits have stopped being twins. The trade deficit grew smaller in the early 1990s as the budget deficit increased, and then the trade deficit grew larger in the late 1990s as the budget deficit turned into a surplus. In the first half of the 2000s, both budget and trade deficits increased. But in 2009, the trade deficit declined as the budget deficit increased.

Of course, no one should expect the budget deficit and trade deficit to move in lockstep, because the other parts of the national saving and investment identity—investment and private savings—will often change as well. In the late 1990s, for example, the government budget balance turned from deficit to surplus, but the trade deficit remained large and growing. During this time, the inflow of foreign financial investment was supporting a surge of physical capital investment by U.S. firms. In the first half of the 2000s, the budget and trade deficits again increased together, but in 2009, the budget deficit increased while the trade deficit declined. The budget deficit and the trade deficits are related to each other, but they are more like cousins than twins.

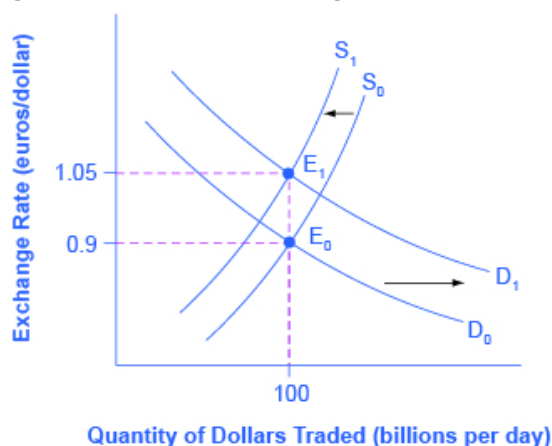
Budget Deficits and Exchange Rates

Exchange rates can also help to explain why budget deficits are linked to trade deficits. [\[link\]](#) shows a situation using the exchange rate for the U.S. dollar, measured in euros. At the original equilibrium (E_0), where the demand for U.S. dollars (D_0) intersects with the supply of U.S. dollars (S_0) on the foreign exchange market, the exchange rate is 0.9 euros per U.S. dollar and the equilibrium quantity traded in the market is \$100 billion per day (which was roughly the quantity of dollar–euro trading in exchange rate markets in the

mid-2000s). Then the U.S. budget deficit rises and foreign financial investment provides the source of funds for that budget deficit.

International financial investors, as a group, will demand more U.S. dollars on foreign exchange markets to purchase the U.S. government bonds, and they will supply fewer of the U.S. dollars that they already hold in these markets. Demand for U.S. dollars on the foreign exchange market shifts from D_0 to D_1 and the supply of U.S. dollars falls from S_0 to S_1 . At the new equilibrium (E_1), the exchange rate has appreciated to 1.05 euros per dollar while, in this example, the quantity of dollars traded remains the same.

Budget Deficits and Exchange Rates



Imagine that the U.S. government increases its borrowing and the funds come from European financial investors. To purchase U.S. government bonds, those European investors will need to demand more U.S. dollars on foreign exchange markets, causing the demand for U.S. dollars to shift to the right from D_0 to D_1 . European financial investors as a group will also be less likely to supply U.S. dollars to the foreign exchange markets, causing the supply of U.S. dollars to shift from S_0 to S_1 . The equilibrium exchange rate strengthens from 0.9 euro/ dollar at E_0 to 1.05 euros/dollar at E_1 .

A stronger exchange rate, of course, makes it more difficult for exporters to sell their goods abroad while making imports cheaper, so a trade deficit (or a reduced trade surplus) results.

Thus, a budget deficit can easily result in an inflow of foreign financial capital, a stronger exchange rate, and a trade deficit.

You can also imagine this appreciation of the exchange rate as being driven by interest rates. As explained earlier in [Budget Deficits and Interest Rates in Fiscal Policy, Investment, and Economic Growth](#), a budget deficit increases demand in markets for domestic financial capital, raising the domestic interest rate. A higher interest rate will attract an inflow of foreign financial capital, and appreciate the exchange rate in response to the increase in demand for U.S. dollars by foreign investors and a decrease in supply of U. S. dollars. Because of higher interest rates in the United States, Americans find U.S. bonds more attractive than foreign bonds. When Americans are buying fewer foreign bonds, they are supplying fewer U.S. dollars. The depreciation of the U.S. dollar leads to a larger trade deficit (or reduced surplus). The connections between inflows of foreign investment capital, interest rates, and exchange rates are all just different ways of drawing the same economic connections: a larger budget deficit can result in a larger trade deficit, although the connection should not be expected to be one-to-one.

From Budget Deficits to International Economic Crisis

The economic story of how an outflow of international financial capital can cause a deep recession is laid out, step-by-step, in the [Exchange Rates and International Capital Flows](#) chapter. When international financial investors decide to withdraw their funds from a country like Turkey, they increase the supply of the Turkish lira and reduce the demand for lira, depreciating the lira exchange rate. When firms and the government in a country like Turkey borrow money in international financial markets, they typically do so in stages. First, banks in Turkey borrow in a widely used currency like U.S. dollars or euros, then convert those U.S. dollars to lira, and then lend the money to borrowers in Turkey. If the value of the lira exchange rate depreciates, then Turkey's banks will find it impossible to repay the international loans that are in U.S. dollars or euros.

The combination of less foreign investment capital and banks that are bankrupt can sharply reduce aggregate demand, which causes a deep recession. Many countries around the world have experienced this kind of recession in recent years: along with Turkey in 2002, this general pattern was followed by Mexico in 1995, Thailand and countries across East Asia in 1997–1998, Russia in 1998, and Argentina in 2002. In many of these countries, large government budget deficits played a role in setting the stage for the financial crisis. A moderate increase in a budget deficit that leads to a moderate increase in a trade deficit and a moderate appreciation of the exchange rate is not necessarily a cause for concern. But beyond some point that is hard to define in advance, a series of large budget deficits can become a cause for concern among international investors.

One reason for concern is that extremely large budget deficits mean that aggregate demand may shift so far to the right as to cause high inflation. The example of Turkey is a situation where very large budget deficits brought inflation rates well into double digits. In addition, very large budget deficits at some point begin to raise a fear that the borrowing will not be

repaid. In the last 175 years, the government of Turkey has been unable to pay its debts and defaulted on its loans six times. Brazil's government has been unable to pay its debts and defaulted on its loans seven times; Venezuela, nine times; and Argentina, five times. The risk of high inflation or a default on repaying international loans will worry international investors, since both factors imply that the rate of return on their investments in that country may end up lower than expected. If international investors start withdrawing the funds from a country rapidly, the scenario of less investment, a depreciated exchange rate, widespread bank failure, and deep recession can occur. The following Clear It Up feature explains other impacts of large deficits.

Note:

What are the risks of chronic large deficits in the United States?

If a government runs large budget deficits for a sustained period of time, what can go wrong? According to a recent report by the Brookings Institution, a key risk of a large budget deficit is that government debt may grow too high compared to the country's GDP growth. As debt grows, the national savings rate will decline, leaving less available in financial capital for private investment. The impact of chronically large budget deficits is as follows:

- As the population ages, there will be an increasing demand for government services that may cause higher government deficits. Government borrowing and its interest payments will pull resources away from domestic investment in human capital and physical capital that is essential to economic growth.
- Interest rates may start to rise so that the cost of financing government debt will rise as well, creating pressure on the government to reduce its budget deficits through spending cuts and tax increases. These steps will be politically painful, and they will also have a contractionary effect on aggregate demand in the economy.
- Rising percentage of debt to GDP will create uncertainty in the financial and global markets that might cause a country to resort to inflationary tactics to reduce the real value of the debt outstanding. This will decrease real wealth and damage confidence in the country's ability to manage its spending. After all, if the government has borrowed at a fixed interest rate of, say, 5%, and it lets inflation rise above that 5%, then it will effectively be able to repay its debt at a negative real interest rate.

The conventional reasoning suggests that the relationship between sustained deficits that lead to high levels of government debt and long-term growth is negative. How significant this relationship is, how big an issue it is compared to other macroeconomic issues, and the direction of causality, is less clear.

What remains important to acknowledge is that the relationship between debt and growth is negative and that for some countries, the relationship may be stronger than in others. It is also important to acknowledge the direction of causality: does high debt cause slow growth, slow growth cause high debt, or are both high debt and slow growth the result of

third factors? In our analysis, we have argued simply that high debt causes slow growth. There may be more to this debate than we have space to discuss here.

Using Fiscal Policy to Address Trade Imbalances

If a nation is experiencing the inflow of foreign investment capital associated with a trade deficit because foreign investors are making long-term direct investments in firms, there may be no substantial reason for concern. After all, many low-income nations around the world would welcome direct investment by multinational firms that ties them more closely into the global networks of production and distribution of goods and services. In this case, the inflows of foreign investment capital and the trade deficit are attracted by the opportunities for a good rate of return on private sector investment in an economy.

However, governments should beware of a sustained pattern of high budget deficits and high trade deficits. The danger arises in particular when the inflow of foreign investment capital is not funding long-term physical capital investment by firms, but instead is short-term portfolio investment in government bonds. When inflows of foreign financial investment reach high levels, foreign financial investors will be on the alert for any reason to fear that the country's exchange rate may decline or the government may be unable to repay what it has borrowed on time. Just as a few falling rocks can trigger an avalanche; a relatively small piece of bad news about an economy can trigger an enormous outflow of short-term financial capital.

Reducing a nation's budget deficit will not always be a successful method of reducing its trade deficit, because other elements of the national saving and investment identity, like private saving or investment, may change instead. In those cases when the budget deficit is the main cause of the trade deficit, governments should take steps to reduce their budget deficits, lest they make their economy vulnerable to a rapid outflow of international financial capital that could bring a deep recession.

Note:

Financing Higher Education

Over the period between 1982 and 2012, the increases in the cost of a college education had far outpaced that of the income of the typical American family. According to the research done by the President Obama's staff, the cost of education at a four-year public college increased by 257% compared to an increase in family incomes of only 16% over the prior 30 years. The ongoing debate over a balanced budget and proposed cutbacks accentuated the need to increase investment in human capital to grow the economy versus deepening the already significant debt levels of the U.S. government. In the summer of 2013, President Obama presented a plan to make college more affordable that included increasing Pell Grant awards and the number of recipients, caps on interest rates for

student loans, and providing education tax credits. In addition, the plan includes an accountability method for institutions of higher education that focuses on completion rates and creates a College Scorecard. Whether or not all these initiatives come to fruition remains to be seen, but they are indicative of creative approaches that government can take to meet its obligation from both a public and fiscal policy perspective.

Key Concepts and Summary

The government need not balance its budget every year. However, a sustained pattern of large budget deficits over time risks causing several negative macroeconomic outcomes: a shift to the right in aggregate demand that causes an inflationary increase in the price level; crowding out private investment in physical capital in a way that slows down economic growth; and creating a dependence on inflows of international portfolio investment which can sometimes turn into outflows of foreign financial investment that can be injurious to a macroeconomy.

Self-Check Questions

Exercise:

Problem:

In the late 1990s, the U.S. government moved from a budget deficit to a budget surplus and the trade deficit in the U.S. economy grew substantially. Using the national saving and investment identity, what can you say about the direction in which saving and/or investment must have changed in this economy?

Solution:

In this case, the national saving and investment identity is written in this way:

Equation:

$$\text{Quantity supplied of financial capital} = \text{Quantity demanded of financial capital} \\ (T - G) + (M - X) + S = I$$

The increase in the government budget surplus and the increase in the trade deficit both increased the supply of financial capital. If investment in physical capital remained unchanged, then private savings must go down, and if savings remained unchanged, then investment must go up. In fact, both effects happened; that is, in the late 1990s, in the U.S. economy, savings declined and investment rose.

Review Questions

Exercise:

Problem: Under what conditions will a larger budget deficit cause a trade deficit?

Critical Thinking Questions

Exercise:

Problem:

Explain how a shift from a government budget deficit to a budget surplus might affect the exchange rate.

Exercise:

Problem:

Describe how a plan for reducing the government deficit might affect a college student, a young professional, and a middle-income family.

Problems

Exercise:

Problem:

Sketch a diagram of how a budget deficit causes a trade deficit. (*Hint:* Begin with what will happen to the exchange rate when foreigners demand more U.S. government debt.)

Exercise:

Problem:

Sketch a diagram of how sustained budget deficits cause low economic growth.

Exercise:

Problem:

Assume that you are employed by the government of Tanzania in 1964, a new nation recently independent from Britain. The Tanzanian parliament has decided that it will spend 10 million shillings on schools, roads, and healthcare for the year. You estimate that the net taxes for the year are eight million shillings. The difference will be financed by selling 10-year government bonds at 12% interest per year. The interest on outstanding bonds must be added to government expenditure each year. Assume that additional taxes are added to finance this increase in government expenditure so the gap between government spending is always two million. If the school, road, and healthcare budget are unchanged, compute the value of the accumulated debt in 10 years.

References

The White House. "This is why it's time to make college more affordable." Last modified August 20, 2013. <http://www.whitehouse.gov/share/college-affordability>.

Rubin, Robert E., Peter R. Orszag, and Allen Sinai. "Sustained Budget Deficits: Longer-Run U.S. Economic Performance and the Risk of Financial and Fiscal Disarray." Last modified January 4, 2004.
<http://www.brookings.edu/~media/research/files/papers/2004/1/05budgetdeficit%20orszag/20040105.pdf>.

Introduction to Money and Banking

class="introduction"

Cowrie Shell or Money?

Is this an image
of a cowrie
shell or money?

The answer is:

Both. For
centuries, the
extremely
durable cowrie
shell was used
as a medium of
exchange in
various parts of
the world.

(Credit:
modification of
work by
“prilfish”/Flick
r Creative
Commons)



Note:

The Many Disguises of Money: From Cowries to Bitcoins

Here is a trivia question: In the history of the world, what item was used for money over the broadest geographic area and for the longest period of time? The answer is not gold, silver, or any precious metal. It is the cowrie, a mollusk shell found mainly off the Maldives Islands in the Indian Ocean. Cowries served as money as early as 700 B.C. in China. By the 1500s, they were in widespread use across India and Africa. For several centuries after that, cowries were used in markets including southern Europe, western Africa, India, and China for a wide range of purchases: everything from buying lunch or a ferry ride to paying for a shipload of silk or rice. Cowries were still acceptable as a way of paying taxes in certain African nations in the early twentieth century.

What made cowries work so well as money? First, they are extremely durable—lasting a century or more. As the late economic historian Karl Polanyi put it, they can be “poured, sacked, shoveled, hoarded in heaps” while remaining “clean, dainty, stainless, polished, and milk-white.”

Second, parties could use cowries either by counting shells of a certain

size, or—for large purchases—by measuring the weight or volume of the total shells to be exchanged. Third, it was impossible to counterfeit a cowrie shell, but gold or silver coins could be counterfeited by making copies with cheaper metals. Finally, in the heyday of cowrie money, from the 1500s into the 1800s, the collection of cowries was tightly controlled, first by the Portuguese and later by the Dutch and the English. As a result, the supply of cowries was allowed to grow quickly enough to serve the needs of commerce, but not so quickly that they were no longer scarce. Money throughout the ages has taken many different forms and continues to evolve even today. What do you think money is?

Note:

Introduction to Money and Banking

In this chapter, you will learn about:

- Defining Money by Its Functions
- Measuring Money: Currency, M1, and M2
- The Role of Banks
- How Banks Create Money

The discussion of money and banking is a central component in the study of macroeconomics. At this point, you should have firmly in mind the main goals of macroeconomics from [Welcome to Economics!](#): economic growth, low unemployment, and low inflation. We have yet to discuss money and its role in helping to achieve our macroeconomic goals.

You should also understand Keynesian and neoclassical frameworks for macroeconomic analysis and how these frameworks can be embodied in the aggregate demand/aggregate supply (AD/AS) model. With the goals and frameworks for macroeconomic analysis in mind, the final step is to discuss the two main categories of macroeconomic policy: monetary policy, which focuses on money, banking and interest rates; and fiscal policy, which focuses on government spending, taxes, and borrowing. This chapter

discusses what economists mean by money, and how money is closely interrelated with the banking system. [Monetary Policy and Bank Regulation](#) furthers this discussion.

Defining Money by Its Functions

By the end of this section, you will be able to:

- Explain the various functions of money
- Contrast commodity money and fiat money

Money for the sake of money is not an end in itself. You cannot eat dollar bills or wear your bank account. Ultimately, the usefulness of money rests in exchanging it for goods or services. As the American writer and humorist Ambrose Bierce (1842–1914) wrote in 1911, money is a “blessing that is of no advantage to us excepting when we part with it.” Money is what people regularly use when purchasing or selling goods and services, and thus money must be widely accepted by both buyers and sellers. This concept of money is intentionally flexible, because money has taken a wide variety of forms in different cultures.

Barter and the Double Coincidence of Wants

To understand the usefulness of money, we must consider what the world would be like without money. How would people exchange goods and services? Economies without money typically engage in the barter system. **Barter**—literally trading one good or service for another—is highly inefficient for trying to coordinate the trades in a modern advanced economy. In an economy without money, an exchange between two people would involve a **double coincidence of wants**, a situation in which two people each want some good or service that the other person can provide. For example, if an accountant wants a pair of shoes, this accountant must find someone who has a pair of shoes in the correct size and who is willing to exchange the shoes for some hours of accounting services. Such a trade is likely to be difficult to arrange. Think about the complexity of such trades in a modern economy, with its extensive division of labor that involves thousands upon thousands of different jobs and goods.

Another problem with the barter system is that it does not allow us to easily enter into future contracts for the purchase of many goods and services. For example, if the goods are perishable it may be difficult to exchange them

for other goods in the future. Imagine a farmer wanting to buy a tractor in six months using a fresh crop of strawberries. Additionally, while the barter system might work adequately in small economies, it will keep these economies from growing. The time that individuals would otherwise spend producing goods and services and enjoying leisure time is spent bartering.

Functions for Money

Money solves the problems created by the barter system. (We will get to its definition soon.) First, money serves as a **medium of exchange**, which means that money acts as an intermediary between the buyer and the seller. Instead of exchanging accounting services for shoes, the accountant now exchanges accounting services for money. This money is then used to buy shoes. To serve as a medium of exchange, money must be very widely accepted as a method of payment in the markets for goods, labor, and financial capital.

Second, money must serve as a **store of value**. In a barter system, we saw the example of the shoemaker trading shoes for accounting services. But she risks having her shoes go out of style, especially if she keeps them in a warehouse for future use—their value will decrease with each season. Shoes are not a good store of value. Holding money is a much easier way of storing value. You know that you do not need to spend it immediately because it will still hold its value the next day, or the next year. This function of money does not require that money is a *perfect* store of value. In an economy with inflation, money loses some buying power each year, but it remains money.

Third, money serves as a **unit of account**, which means that it is the ruler by which other values are measured. For example, an accountant may charge \$100 to file your tax return. That \$100 can purchase two pair of shoes at \$50 a pair. Money acts as a common denominator, an accounting method that simplifies thinking about trade-offs.

Finally, another function of money is that money must serve as a **standard of deferred payment**. This means that if money is usable today to make purchases, it must also be acceptable to make purchases today that will be

paid in the *future*. Loans and future agreements are stated in monetary terms and the standard of deferred payment is what allows us to buy goods and services today and pay in the future. So **money** serves all of these functions — it is a medium of exchange, store of value, unit of account, and standard of deferred payment.

Commodity versus Fiat Money

Money has taken a wide variety of forms in different cultures. Gold, silver, cowrie shells, cigarettes, and even cocoa beans have been used as money. Although these items are used as **commodity money**, they also have a value from use as something other than money. Gold, for example, has been used throughout the ages as money although today it is not used as money but rather is valued for its other attributes. Gold is a good conductor of electricity and is used in the electronics and aerospace industry. Gold is also used in the manufacturing of energy efficient reflective glass for skyscrapers and is used in the medical industry as well. Of course, gold also has value because of its beauty and malleability in the creation of jewelry.

As commodity money, gold has historically served its purpose as a medium of exchange, a store of value, and as a unit of account. **Commodity-backed currencies** are dollar bills or other currencies with values backed up by gold or other commodity held at a bank. During much of its history, the money supply in the United States was backed by gold and silver. Interestingly, antique dollars dated as late as 1957, have “Silver Certificate” printed over the portrait of George Washington, as shown in [\[link\]](#). This meant that the holder could take the bill to the appropriate bank and exchange it for a dollar’s worth of silver.

A Silver Certificate and a Modern U.S. Bill



Until 1958, silver certificates were commodity-backed money—backed by silver, as indicated by the words “Silver Certificate” printed on the bill. Today, U.S. bills are backed by the Federal Reserve, but as fiat money. (Credit: “The.Comedian”/Flickr Creative Commons)

As economies grew and became more global in nature, the use of commodity monies became more cumbersome. Countries moved towards the use of **fiat money**. Fiat money has no intrinsic value, but is declared by a government to be the legal tender of a country. The United States’ paper money, for example, carries the statement: “THIS NOTE IS LEGAL TENDER FOR ALL DEBTS, PUBLIC AND PRIVATE.” In other words, by government decree, if you owe a debt, then legally speaking, you can pay that debt with the U.S. currency, even though it is not backed by a commodity. The only backing of our money is universal faith and trust that the currency has value, and nothing more.

Note:

Watch this [video](#) on the “History of Money.”



Key Concepts and Summary

Money is what people in a society regularly use when purchasing or selling goods and services. If money were not available, people would need to barter with each other, meaning that each person would need to identify others with whom they have a double coincidence of wants—that is, each party has a specific good or service that the other desires. Money serves several functions: a medium of exchange, a unit of account, a store of value, and a standard of deferred payment. There are two types of money: commodity money, which is an item used as money, but which also has value from its use as something other than money; and fiat money, which has no intrinsic value, but is declared by a government to be the legal tender of a country.

Self-Check Questions

Exercise:**Problem:**

In many casinos, a person buys chips to use for gambling. Within the walls of the casino, these chips can often be used to buy food and drink or even a hotel room. Do chips in a gambling casino serve all three functions of money?

Solution:

As long as you remain within the walls of the casino, chips fit the definition of money; that is, they serve as a medium of exchange, a unit of account, and a store of value. Chips do not work very well as money once you leave the casino, but many kinds of money do not work well in other areas. For example, it is hard to spend money from Turkey or Brazil at your local supermarket or at the movie theater.

Exercise:**Problem:**

Can you name some item that is a store of value, but does not serve the other functions of money?

Solution:

Many physical items that a person buys at one time but may sell at another time can serve as an answer to this question. Examples include a house, land, art, rare coins or stamps, and so on.

Review Questions**Exercise:**

Problem: What are the four functions served by money?

Exercise:**Problem:**

How does the existence of money simplify the process of buying and selling?

Exercise:

Problem: What is the double-coincidence of wants?

Critical Thinking Questions

Exercise:

Problem:

The Bring it Home Feature discusses the use of cowrie shells as money. Although cowrie shells are no longer used as money, do you think other forms of commodity monies are possible? What role might technology play in our definition of money?

Exercise:

Problem:

Imagine that you are a barber in a world without money. Explain why it would be tricky to obtain groceries, clothing, and a place to live.

References

Hogendorn, Jan and Marion Johnson. *The Shell Money of the Slave Trade*. Cambridge University Press, 2003. 6.

Glossary

barter

literally, trading one good or service for another, without using money

commodity money

an item that is used as money, but which also has value from its use as something other than money

commodity-backed currencies

are dollar bills or other currencies with values backed up by gold or another commodity

double coincidence of wants

a situation in which two people each want some good or service that the other person can provide

fiat money

has no intrinsic value, but is declared by a government to be the legal tender of a country

medium of exchange

whatever is widely accepted as a method of payment

money

whatever serves society in four functions: as a medium of exchange, a store of value, a unit of account, and a standard of deferred payment.

standard of deferred payment

money must also be acceptable to make purchases today that will be paid in the future

store of value

something that serves as a way of preserving economic value that can be spent or consumed in the future

unit of account

the common way in which market values are measured in an economy

Measuring Money: Currency, M1, and M2

By the end of this section, you will be able to:

- Contrast M1 money supply and M2 money supply
- Classify monies as M1 money supply or M2 money supply

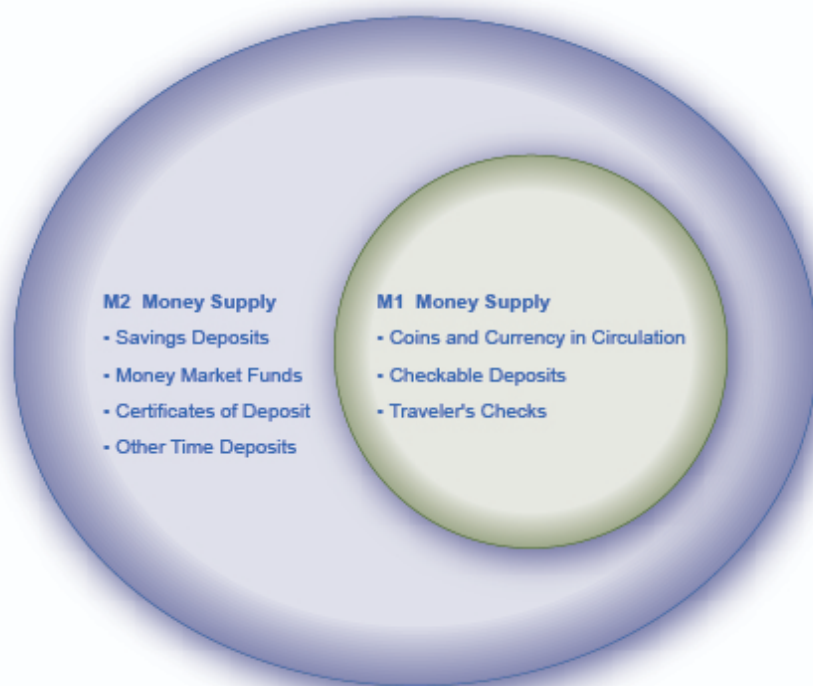
Cash in your pocket certainly serves as money. But what about checks or credit cards? Are they money, too? Rather than trying to state a single way of measuring money, economists offer broader definitions of money based on liquidity. Liquidity refers to how quickly a financial asset can be used to buy a good or service. For example, cash is very liquid. Your \$10 bill can be easily used to buy a hamburger at lunchtime. However, \$10 that you have in your savings account is not so easy to use. You must go to the bank or ATM machine and withdraw that cash to buy your lunch. Thus, \$10 in your savings account is *less* liquid.

The Federal Reserve Bank, which is the central bank of the United States, is a bank regulator and is responsible for monetary policy and defines money according to its liquidity. There are two definitions of money: M1 and M2 money supply. **M1 money supply** includes those monies that are very liquid such as cash, checkable (demand) deposits, and traveler's checks. **M2 money supply** is less liquid in nature and includes M1 plus savings and time deposits, certificates of deposits, and money market funds.

M1 money supply includes **coins and currency in circulation**—the coins and bills that circulate in an economy that are not held by the U.S. Treasury, at the Federal Reserve Bank, or in bank vaults. Closely related to currency are checkable deposits, also known as **demand deposits**. These are the amounts held in checking accounts. They are called demand deposits or checkable deposits because the banking institution must give the deposit holder his money “on demand” when a check is written or a debit card is used. These items together—currency, and checking accounts in banks—make up the definition of money known as M1, which is measured daily by the Federal Reserve System. Traveler's checks are also included in M1, but have decreased in use over the recent past.

A broader definition of money, M2 includes everything in M1 but also adds other types of deposits. For example, M2 includes **savings deposits** in banks, which are bank accounts on which you cannot write a check directly, but from which you can easily withdraw the money at an automatic teller machine or bank. Many banks and other financial institutions also offer a chance to invest in **money market funds**, where the deposits of many individual investors are pooled together and invested in a safe way, such as short-term government bonds. Another ingredient of M2 are the relatively small (that is, less than about \$100,000) certificates of deposit (CDs) or **time deposits**, which are accounts that the depositor has committed to leaving in the bank for a certain period of time, ranging from a few months to a few years, in exchange for a higher interest rate. In short, all these types of M2 are money that you can withdraw and spend, but which require a greater effort to do so than the items in M1 [\[link\]](#) should help in visualizing the relationship between M1 and M2. Note that M1 is included in the M2 calculation.

The Relationship between M1 and M2 Money



M1 and M2 money have several definitions, ranging from narrow to broad. M1 = coins and currency in circulation + checkable (demand) deposit + traveler's

checks. $M2 = M1 + \text{savings deposits} + \text{money market funds} + \text{certificates of deposit} + \text{other time deposits}$.

The Federal Reserve System is responsible for tracking the amounts of M1 and M2 and prepares a weekly release of information about the money supply. To provide an idea of what these amounts sound like, according to the Federal Reserve Bank's measure of the U.S. money stock, at the end of February 2015, M1 in the United States was \$3 trillion, while M2 was \$11.8 trillion. A breakdown of the portion of each type of money that comprised M1 and M2 in February 2015, as provided by the Federal Reserve Bank, is provided in [\[link\]](#).

Components of M1 in the U.S. (February 2015, Seasonally Adjusted)	\$ billions
Currency	\$1,271.8
Traveler's checks	\$2.9
Demand deposits and other checking accounts	\$1,713.5
<i>Total M1</i>	<i>\$2,988.2 (or \$3 trillion)</i>
Components of M2 in the U.S. (February 2015, Seasonally Adjusted)	\$ billions
M1 money supply	\$2,988.2

Savings accounts	\$7,712.1
Time deposits	\$509.2
Individual money market mutual fund balances	\$610.8
<i>Total M2</i>	<i>\$11,820.3 (or \$11.8 trillion)</i>

M1 and M2 Federal Reserve Statistical Release, Money Stock Measures (Source: Federal Reserve Statistical Release, <http://www.federalreserve.gov/RELEASES/h6/current/default.htm#t2tg1link>)

The lines separating M1 and M2 can become a little blurry. Sometimes elements of M1 are not treated alike; for example, some businesses will not accept personal checks for large amounts, but will accept traveler's checks or cash. Changes in banking practices and technology have made the savings accounts in M2 more similar to the checking accounts in M1. For example, some savings accounts will allow depositors to write checks, use automatic teller machines, and pay bills over the Internet, which has made it easier to access savings accounts. As with many other economic terms and statistics, the important point is to know the strengths and limitations of the various definitions of money, not to believe that such definitions are as clear-cut to economists as, say, the definition of nitrogen is to chemists.

Where does "plastic money" like debit cards, credit cards, and smart money fit into this picture? A **debit card**, like a check, is an instruction to the user's bank to transfer money directly and immediately from your bank account to the seller. It is important to note that in our definition of money, it is *checkable deposits* that are money, not the paper check or the debit card. Although you can make a purchase with a **credit card**, it is not considered money but rather a short term loan from the credit card company to you. When you make a purchase with a credit card, the credit card company immediately transfers money from its checking account to the seller, and at the end of the month, the credit card company sends you a bill for what you have charged that month. Until you pay the credit card bill, you have effectively borrowed money from the credit card company. With a **smart**

card, you can store a certain value of money on the card and then use the card to make purchases. Some “smart cards” used for specific purposes, like long-distance phone calls or making purchases at a campus bookstore and cafeteria, are not really all that smart, because they can only be used for certain purchases or in certain places.

In short, credit cards, debit cards, and smart cards are different ways to move money when a purchase is made. But having more credit cards or debit cards does not change the quantity of money in the economy, any more than having more checks printed increases the amount of money in your checking account.

One key message underlying this discussion of M1 and M2 is that money in a modern economy is not just paper bills and coins; instead, money is closely linked to bank accounts. Indeed, the macroeconomic policies concerning money are largely conducted through the banking system. The next section explains how banks function and how a nation’s banking system has the power to create money.

Note:

Read a brief [article](#) on the current monetary challenges in Sweden.



Key Concepts and Summary

Money is measured with several definitions: M1 includes currency and money in checking accounts (demand deposits). Traveler’s checks are also a component of M1, but are declining in use. M2 includes all of M1, plus

savings deposits, time deposits like certificates of deposit, and money market funds.

Self-Check Questions

Exercise:

Problem:

If you are out shopping for clothes and books, what is easiest and most convenient for you to spend: M1 or M2? Explain your answer.

Solution:

The currency and checks in M1 are easiest to spend. It is harder to spend M2 directly, although if there is an automatic teller machine in the shopping mall, you can turn M2 from your savings account into an M1 of currency quite quickly. If your answer is about “credit cards,” then you are really talking about spending M1—although it is M1 from the account of the credit card company, which you will repay later when your credit card bill comes due.

Exercise:

Problem:

For the following list of items, indicate if they are in M1, M2, or neither:

- a. Your \$5,000 line of credit on your Bank of America card
 - b. \$50 dollars' worth of traveler's checks you have not used yet
 - c. \$1 in quarters in your pocket
 - d. \$1200 in your checking account
 - e. \$2000 you have in a money market account
-

Solution:

- a. Neither in M1 or M2

- b. That is part of M1, and because M2 includes M1 it is also part of M2
- c. Currency out in the public hands is part of M1 and M2
- d. Checking deposits are in M1 and M2
- e. Money market accounts are in M2

Review Questions

Exercise:

Problem: What components of money are counted as part of M1?

Exercise:

Problem: What components of money are counted in M2?

Critical Thinking Questions

Exercise:

Problem:

Explain why you think the Federal Reserve Bank tracks M1 and M2.

Exercise:

Problem:

The total amount of U.S. currency in circulation divided by the U.S. population comes out to about \$3,500 per person. That is more than most of us carry. Where is all the cash?

Exercise:

Problem:

If you take \$100 out of your piggy bank and deposit it in your checking account, how did M1 change? Did M2 change?

References

Federal Reserve Statistical Release. November 23, 2013.

<http://www.federalreserve.gov/RELEASES/h6/current/default.htm#t2tg1link>

.

Glossary

coins and currency in circulation

the coins and bills that circulate in an economy that are not held by the U.S Treasury, at the Federal Reserve Bank, or in bank vaults

credit card

immediately transfers money from the credit card company's checking account to the seller, and at the end of the month the user owes the money to the credit card company; a credit card is a short-term loan

debit card

like a check, is an instruction to the user's bank to transfer money directly and immediately from your bank account to the seller

demand deposit

checkable deposit in banks that is available by making a cash withdrawal or writing a check

M1 money supply

a narrow definition of the money supply that includes currency and checking accounts in banks, and to a lesser degree, traveler's checks.

M2 money supply

a definition of the money supply that includes everything in M1, but also adds savings deposits, money market funds, and certificates of deposit

money market fund

the deposits of many investors are pooled together and invested in a safe way like short-term government bonds

savings deposit

bank account where you cannot withdraw money by writing a check, but can withdraw the money at a bank—or can transfer it easily to a checking account

smart card

stores a certain value of money on a card and then the card can be used to make purchases

time deposit

account that the depositor has committed to leaving in the bank for a certain period of time, in exchange for a higher rate of interest; also called certificate of deposit

The Role of Banks

By the end of this section, you will be able to:

- Explain how banks act as intermediaries between savers and borrowers
- Evaluate the relationship between banks, savings and loans, and credit unions
- Analyze the causes of bankruptcy and recessions

The late bank robber named Willie Sutton was once asked why he robbed banks. He answered: “That’s where the money is.” While this may have been true at one time, from the perspective of modern economists, Sutton is both right and wrong. He is wrong because the overwhelming majority of money in the economy is not in the form of currency sitting in vaults or drawers at banks, waiting for a robber to appear. Most money is in the form of bank accounts, which exist only as electronic records on computers. From a broader perspective, however, the bank robber was more right than he may have known. Banking is intimately interconnected with money and consequently, with the broader economy.

Banks make it far easier for a complex economy to carry out the extraordinary range of transactions that occur in goods, labor, and financial capital markets. Imagine for a moment what the economy would be like if all payments had to be made in cash. When shopping for a large purchase or going on vacation you might need to carry hundreds of dollars in a pocket or purse. Even small businesses would need stockpiles of cash to pay workers and to purchase supplies. A bank allows people and businesses to store this money in either a checking account or savings account, for example, and then withdraw this money as needed through the use of a direct withdrawal, writing a check, or using a debit card.

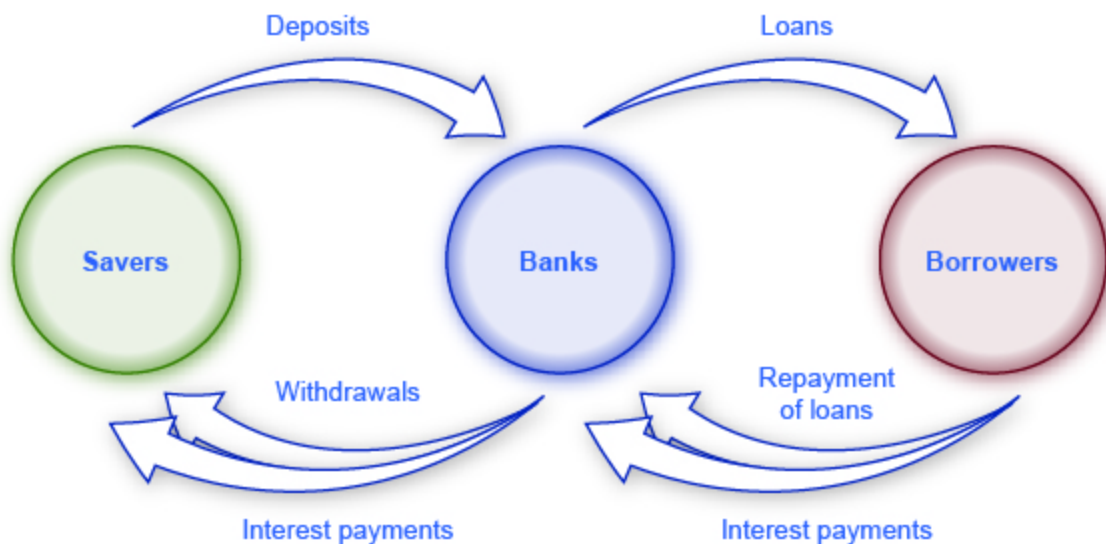
Banks are a critical intermediary in what is called the **payment system**, which helps an economy exchange goods and services for money or other financial assets. Also, those with extra money that they would like to save can store their money in a bank rather than look for an individual that is willing to borrow it from them and then repay them at a later date. Those who want to borrow money can go directly to a bank rather than trying to find someone to lend them cash **Transaction costs** are the costs associated

with finding a lender or a borrower for this money. Thus, banks lower transactions costs and act as financial intermediaries—they bring savers and borrowers together. Along with making transactions much safer and easier, banks also play a key role in the creation of money.

Banks as Financial Intermediaries

An “intermediary” is one who stands between two other parties. Banks are a **financial intermediary**—that is, an institution that operates between a saver who deposits money in a bank and a borrower who receives a loan from that bank. Financial intermediaries include other institutions in the financial market such as insurance companies and pension funds, but they will not be included in this discussion because they are not considered to be **depository institutions**, which are institutions that accept money *deposits* and then use these to make loans. All the funds deposited are mingled in one big pool, which is then loaned out. [\[link\]](#) illustrates the position of banks as financial intermediaries, with deposits flowing into a bank and loans flowing out. Of course, when banks make loans to firms, the banks will try to funnel financial capital to healthy businesses that have good prospects for repaying the loans, not to firms that are suffering losses and may be unable to repay.

Banks as Financial Intermediaries



Banks act as financial intermediaries because they stand between

savers and borrowers. Savers place deposits with banks, and then receive interest payments and withdraw money. Borrowers receive loans from banks and repay the loans with interest. In turn, banks return money to savers in the form of withdrawals, which also include interest payments from banks to savers.

Note:

How are banks, savings and loans, and credit unions related?

Banks have a couple of close cousins: savings institutions and credit unions. Banks, as explained, receive deposits from individuals and businesses and make loans with the money. Savings institutions are also sometimes called “savings and loans” or “thrifts.” They also take loans and make deposits. However, from the 1930s until the 1980s, federal law limited how much interest savings institutions were allowed to pay to depositors. They were also required to make most of their loans in the form of housing-related loans, either to homebuyers or to real-estate developers and builders.

A credit union is a nonprofit financial institution that its members own and run. Members of each credit union decide who is eligible to be a member. Usually, potential members would be everyone in a certain community, or groups of employees, or members of a certain organization. The credit union accepts deposits from members and focuses on making loans back to its members. While there are more credit unions than banks and more banks than savings and loans, the total assets of credit unions are growing. In 2008, there were 7,085 banks. Due to the bank failures of 2007–2009 and bank mergers, there were 5,571 banks in the United States at the end of the fourth quarter in 2014. According to the Credit Union National Association, as of December 2014 there were 6,535 credit unions with assets totaling \$1.1 billion. A day of “Transfer Your Money” took place in 2009 out of general public disgust with big bank bailouts. People were encouraged to transfer their deposits to credit unions. This has grown into the ongoing Move Your Money Project. Consequently, some now hold deposits as large as \$50 billion. However, as of 2013, the 12 largest banks

(0.2%) controlled 69 percent of all banking assets, according to the Dallas Federal Reserve.

A Bank's Balance Sheet

A **balance sheet** is an accounting tool that lists assets and liabilities. An **asset** is something of value that is owned and can be used to produce something. For example, the cash you own can be used to pay your tuition. If you own a home, this is also considered an asset. A **liability** is a debt or something you owe. Many people borrow money to buy homes. In this case, a home is the asset, but the mortgage is the liability. The **net worth** is the asset value minus how much is owed (the liability). A bank's balance sheet operates in much the same way. A bank's net worth is also referred to as **bank capital**. A bank has assets such as cash held in its vaults, monies that the bank holds at the Federal Reserve bank (called "reserves"), loans that are made to customers, and bonds.

[\[link\]](#) illustrates a hypothetical and simplified balance sheet for the Safe and Secure Bank. Because of the two-column format of the balance sheet, with the T-shape formed by the vertical line down the middle and the horizontal line under "Assets" and "Liabilities," it is sometimes called a **T-account**.

A Balance Sheet for the Safe and Secure Bank

Assets		Liabilities + Net Worth	
Loans	\$5 million	Deposits	\$10 million
U.S. Government Securities (USGS)	\$4 million		
Reserves	\$2 million	Net Worth	\$1 million

The "T" in a T-account separates the assets of a firm, on the left, from its liabilities, on the right. All firms use T-accounts, though most are much more complex. For a bank, the assets are the financial instruments that either the bank is holding (its reserves) or those instruments where other parties owe money to the bank—like loans made by the bank and U.S. Government Securities, such as U.S. treasury bonds purchased by the bank. Liabilities are what the bank owes to others. Specifically, the bank owes any deposits made in the bank to those who have made them. The net worth of the bank is the total assets minus total liabilities. Net worth is included

on the liabilities side to have the T account balance to zero. For a healthy business, net worth will be positive. For a bankrupt firm, net worth will be negative. In either case, on a bank's T-account, assets will always equal liabilities plus net worth.

When bank customers deposit money into a checking account, savings account, or a certificate of deposit, the bank views these deposits as liabilities. After all, the bank owes these deposits to its customers, when the customers wish to withdraw their money. In the example shown in [\[link\]](#), the Safe and Secure Bank holds \$10 million in deposits.

Loans are the first category of bank assets shown in [\[link\]](#). Say that a family takes out a 30-year mortgage loan to purchase a house, which means that the borrower will repay the loan over the next 30 years. This loan is clearly an asset from the bank's perspective, because the borrower has a legal obligation to make payments to the bank over time. But in practical terms, how can the value of the mortgage loan that is being paid over 30 years be measured in the present? One way of measuring the value of something—whether a loan or anything else—is by estimating what another party in the market is willing to pay for it. Many banks issue home loans, and charge various handling and processing fees for doing so, but then sell the loans to other banks or financial institutions who collect the loan payments. The market where loans are made to borrowers is called the primary loan market, while the market in which these loans are bought and sold by financial institutions is the secondary loan market.

One key factor that affects what financial institutions are willing to pay for a loan, when they buy it in the secondary loan market, is the perceived riskiness of the loan: that is, given the characteristics of the borrower, such as income level and whether the local economy is performing strongly, what proportion of loans of this type will be repaid? The greater the risk that a loan will not be repaid, the less that any financial institution will pay to acquire the loan. Another key factor is to compare the interest rate charged on the original loan with the current interest rate in the economy. If the original loan made at some point in the past requires the borrower to pay a low interest rate, but current interest rates are relatively high, then a financial institution will pay less to acquire the loan. In contrast, if the

original loan requires the borrower to pay a high interest rate, while current interest rates are relatively low, then a financial institution will pay more to acquire the loan. For the Safe and Secure Bank in this example, the total value of its loans if they were sold to other financial institutions in the secondary market is \$5 million.

The second category of bank asset is bonds, which are a common mechanism for borrowing, used by the federal and local government, and also private companies, and nonprofit organizations. A bank takes some of the money it has received in deposits and uses the money to buy bonds—typically bonds issued by the U.S. government. Government bonds are low-risk because the government is virtually certain to pay off the bond, albeit at a low rate of interest. These bonds are an asset for banks in the same way that loans are an asset: The bank will receive a stream of payments in the future. In our example, the Safe and Secure Bank holds bonds worth a total value of \$4 million.

The final entry under assets is **reserves**, which is money that the bank keeps on hand, and that is not loaned out or invested in bonds—and thus does not lead to interest payments. The Federal Reserve requires that banks keep a certain percentage of depositors' money on “reserve,” which means either in their vaults or kept at the Federal Reserve Bank. This is called a reserve requirement. ([Monetary Policy and Bank Regulation](#) will explain how the level of these required reserves are one policy tool that governments have to influence bank behavior.) Additionally, banks may also want to keep a certain amount of reserves on hand in excess of what is required. The Safe and Secure Bank is holding \$2 million in reserves.

The net worth of a bank is defined as its total assets minus its total liabilities. For the Safe and Secure Bank shown in [\[link\]](#), net worth is equal to \$1 million; that is, \$11 million in assets minus \$10 million in liabilities. For a financially healthy bank, the net worth will be positive. If a bank has negative net worth and depositors tried to withdraw their money, the bank would not be able to give all depositors their money.

Note:

For some concrete examples of what banks do, watch this [video](#) from Paul Solman’s “Making Sense of Financial News.”



How Banks Go Bankrupt

A bank that is bankrupt will have a negative net worth, meaning its assets will be worth less than its liabilities. How can this happen? Again, looking at the balance sheet helps to explain.

A well-run bank will assume that a small percentage of borrowers will not repay their loans on time, or at all, and factor these missing payments into its planning. Remember, the calculations of the expenses of banks every year includes a factor for loans that are not repaid, and the value of a bank’s loans on its balance sheet assumes a certain level of riskiness because some loans will not be repaid. Even if a bank expects a certain number of loan defaults, it will suffer if the number of loan defaults is much greater than expected, as can happen during a recession. For example, if the Safe and Secure Bank in [\[link\]](#) experienced a wave of unexpected defaults, so that its loans declined in value from \$5 million to \$3 million, then the assets of the Safe and Secure Bank would decline so that the bank had negative net worth.

Note:

What led to the financial crisis of 2008–2009?

Many banks make mortgage loans so that people can buy a home, but then do not keep the loans on their books as an asset. Instead, the bank sells the

loan. These loans are “securitized,” which means that they are bundled together into a financial security that is sold to investors. Investors in these mortgage-backed securities receive a rate of return based on the level of payments that people make on all the mortgages that stand behind the security.

Securitization offers certain advantages. If a bank makes most of its loans in a local area, then the bank may be financially vulnerable if the local economy declines, so that many people are unable to make their payments. But if a bank sells its local loans, and then buys a mortgage-backed security based on home loans in many parts of the country, it can avoid being exposed to local financial risks. (In the simple example in the text, banks just own “bonds.” In reality, banks can own a number of financial instruments, as long as these financial investments are safe enough to satisfy the government bank regulators.) From the standpoint of a local homebuyer, securitization offers the benefit that a local bank does not need to have lots of extra funds to make a loan, because the bank is only planning to hold that loan for a short time, before selling the loan so that it can be pooled into a financial security.

But securitization also offers one potentially large disadvantage. If a bank is going to hold a mortgage loan as an asset, the bank has an incentive to scrutinize the borrower carefully to ensure that the loan is likely to be repaid. However, a bank that is going to sell the loan may be less careful in making the loan in the first place. The bank will be more willing to make what are called “subprime loans,” which are loans that have characteristics like low or zero down-payment, little scrutiny of whether the borrower has a reliable income, and sometimes low payments for the first year or two that will be followed by much higher payments after that. Some subprime loans made in the mid-2000s were later dubbed NINJA loans: loans made even though the borrower had demonstrated No Income, No Job, or Assets. These subprime loans were typically sold and turned into financial securities—but with a twist. The idea was that if losses occurred on these mortgage-backed securities, certain investors would agree to take the first, say, 5% of such losses. Other investors would agree to take, say, the next 5% of losses. By this approach, still other investors would not need to take any losses unless these mortgage-backed financial securities lost 25% or 30% or more of their total value. These complex securities, along with

other economic factors, encouraged a large expansion of subprime loans in the mid-2000s.

The economic stage was now set for a banking crisis. Banks thought they were buying only ultra-safe securities, because even though the securities were ultimately backed by risky subprime mortgages, the banks only invested in the part of those securities where they were protected from small or moderate levels of losses. But as housing prices fell after 2007, and the deepening recession made it harder for many people to make their mortgage payments, many banks found that their mortgage-backed financial assets could end up being worth much less than they had expected—and so the banks were staring bankruptcy in the face. In the 2008–2011 period, 318 banks failed in the United States.

The risk of an unexpectedly high level of loan defaults can be especially difficult for banks because a bank's liabilities, namely the deposits of its customers, can be withdrawn quickly, but many of the bank's assets like loans and bonds will only be repaid over years or even decades. This **asset-liability time mismatch**—a bank's liabilities can be withdrawn in the short term while its assets are repaid in the long term—can cause severe problems for a bank. For example, imagine a bank that has loaned a substantial amount of money at a certain interest rate, but then sees interest rates rise substantially. The bank can find itself in a precarious situation. If it does not raise the interest rate it pays to depositors, then deposits will flow to other institutions that offer the higher interest rates that are now prevailing. However, if the bank raises the interest rates that it pays to depositors, it may end up in a situation where it is paying a higher interest rate to depositors than it is collecting from those past loans that were made at lower interest rates. Clearly, the bank cannot survive in the long term if it is paying out more in interest to depositors than it is receiving from borrowers.

How can banks protect themselves against an unexpectedly high rate of loan defaults and against the risk of an asset-liability time mismatch? One strategy is for a bank to **diversify** its loans, which means lending to a variety of customers. For example, suppose a bank specialized in lending to

a niche market—say, making a high proportion of its loans to construction companies that build offices in one downtown area. If that one area suffers an unexpected economic downturn, the bank will suffer large losses. However, if a bank loans both to consumers who are buying homes and cars and also to a wide range of firms in many industries and geographic areas, the bank is less exposed to risk. When a bank diversifies its loans, those categories of borrowers who have an unexpectedly large number of defaults will tend to be balanced out, according to random chance, by other borrowers who have an unexpectedly low number of defaults. Thus, diversification of loans can help banks to keep a positive net worth. However, if a widespread recession occurs that touches many industries and geographic areas, diversification will not help.

Along with diversifying their loans, banks have several other strategies to reduce the risk of an unexpectedly large number of loan defaults. For example, banks can sell some of the loans they make in the secondary loan market, as described earlier, and instead hold a greater share of assets in the form of government bonds or reserves. Nevertheless, in a lengthy recession, most banks will see their net worth decline because a higher share of loans will not be repaid in tough economic times.

Key Concepts and Summary

Banks facilitate the use of money for transactions in the economy because people and firms can use bank accounts when selling or buying goods and services, when paying a worker or being paid, and when saving money or receiving a loan. In the financial capital market, banks are financial intermediaries; that is, they operate between savers who supply financial capital and borrowers who demand loans. A balance sheet (sometimes called a T-account) is an accounting tool which lists assets in one column and liabilities in another column. The liabilities of a bank are its deposits. The assets of a bank include its loans, its ownership of bonds, and its reserves (which are not loaned out). The net worth of a bank is calculated by subtracting the bank's liabilities from its assets. Banks run a risk of negative net worth if the value of their assets declines. The value of assets can decline because of an unexpectedly high number of defaults on loans, or if interest rates rise and the bank suffers an asset-liability time mismatch

in which the bank is receiving a low rate of interest on its long-term loans but must pay the currently higher market rate of interest to attract depositors. Banks can protect themselves against these risks by choosing to diversify their loans or to hold a greater proportion of their assets in bonds and reserves. If banks hold only a fraction of their deposits as reserves, then the process of banks' lending money, those loans being re-deposited in banks, and the banks making additional loans will create money in the economy.

Self-Check Questions

Exercise:

Problem:

Explain why the money listed under assets on a bank balance sheet may not actually be in the bank?

Solution:

A bank's assets include cash held in their vaults, but assets also include monies that the bank holds at the Federal Reserve Bank (called "reserves"), loans that are made to customers, and bonds.

Review Questions

Exercise:

Problem: Why is a bank called a financial intermediary?

Exercise:

Problem: What does a balance sheet show?

Exercise:

Problem: What are the assets of a bank? What are its liabilities?

Exercise:

Problem: How do you calculate the net worth of a bank?

Exercise:

Problem: How can a bank end up with negative net worth?

Exercise:

Problem: What is the asset-liability time mismatch that all banks face?

Exercise:

Problem: What is the risk if a bank does not diversify its loans?

Critical Thinking Questions

Exercise:

Problem:

Explain the difference between how you would characterize bank deposits and loans as assets and liabilities on your own personal balance sheet and how a bank would characterize deposits and loans as assets and liabilities on its balance sheet.

Problems

Exercise:

Problem:

A bank has deposits of \$400. It holds reserves of \$50. It has purchased government bonds worth \$70. It has made loans of \$500. Set up a T-account balance sheet for the bank, with assets and liabilities, and calculate the bank's net worth.

References

Credit Union National Association. 2014. "Monthly Credit Union Estimates." Last accessed March 4, 2015. <http://www.cuna.org/Research-And-Strategy/Credit-Union-Data-And-Statistics/>.

Dallas Federal Reserve. 2013. "Ending 'Too Big To Fail': A Proposal for Reform Before It's Too Late". Accessed March 4, 2015. <http://www.dallasfed.org/news/speeches/fisher/2013/fs130116.cfm>.

Richard W. Fisher. "Ending 'Too Big to Fail': A Proposal for Reform Before It's Too Late (With Reference to Patrick Henry, Complexity and Reality) Remarks before the Committee for the Republic, Washington, D.C. Dallas Federal Reserve. January 16, 2013.

"Commercial Banks in the U.S." Federal Reserve Bank of St. Louis. Accessed November 2013. <http://research.stlouisfed.org/fred2/series/USNUM>.

Glossary

asset

item of value owned by a firm or an individual

asset–liability time mismatch

a bank's liabilities can be withdrawn in the short term while its assets are repaid in the long term

balance sheet

an accounting tool that lists assets and liabilities

bank capital

a bank's net worth

depository institution

institution that accepts money deposits and then uses these to make loans

diversify

making loans or investments with a variety of firms, to reduce the risk of being adversely affected by events at one or a few firms

financial intermediary

an institution that operates between a saver with financial assets to invest and an entity who will borrow those assets and pay a rate of return

liability

any amount or debt owed by a firm or an individual

net worth

the excess of the asset value over and above the amount of the liability;
total assets minus total liabilities

payment system

helps an economy exchange goods and services for money or other financial assets

reserves

funds that a bank keeps on hand and that are not loaned out or invested in bonds

T-account

a balance sheet with a two-column format, with the T-shape formed by the vertical line down the middle and the horizontal line under the column headings for “Assets” and “Liabilities”

transaction costs

the costs associated with finding a lender or a borrower for money

How Banks Create Money

By the end of this section, you will be able to:

- Utilize the money multiplier formula to determine how banks create money
- Analyze and create T-account balance sheets
- Evaluate the risks and benefits of money and banks

Banks and money are intertwined. It is not just that most money is in the form of bank accounts. The banking system can literally create money through the process of making loans. Let's see how.

Money Creation by a Single Bank

Start with a hypothetical bank called Singleton Bank. The bank has \$10 million in deposits. The T-account balance sheet for Singleton Bank, when it holds all of the deposits in its vaults, is shown in [\[link\]](#). At this stage, Singleton Bank is simply storing money for depositors and is using these deposits to make loans. In this simplified example, Singleton Bank cannot earn any interest income from these loans and cannot pay its depositors an interest rate either.

Singleton Bank's Balance Sheet: Receives \$10 million in Deposits

Assets		Liabilities + Net Worth	
Reserves	\$10 million	Deposits	\$10 million

Singleton Bank is required by the Federal Reserve to keep \$1 million on reserve (10% of total deposits). It will loan out the remaining \$9 million. By loaning out the \$9 million and charging interest, it will be able to make interest payments to depositors and earn interest income for Singleton Bank (for now, we will keep it simple and not put interest income on the balance sheet). Instead of becoming just a storage place for deposits, Singleton Bank can become a financial intermediary between savers and borrowers.

This change in business plan alters Singleton Bank's balance sheet, as shown in [\[link\]](#). Singleton's assets have changed; it now has \$1 million in reserves and a loan to Hank's Auto Supply of \$9 million. The bank still has \$10 million in deposits.

Singleton Bank's Balance Sheet: 10% Reserves, One Round of Loans

Assets		Liabilities + Net Worth	
Reserves	\$1 million	Deposits	\$10 million
Loan to Hank's Auto Supply	\$9 million		

Singleton Bank lends \$9 million to Hank's Auto Supply. The bank records this loan by making an entry on the balance sheet to indicate that a loan has been made. This loan is an asset, because it will generate interest income for the bank. Of course, the loan officer is not going to let Hank walk out of the bank with \$9 million in cash. The bank issues Hank's Auto Supply a cashier's check for the \$9 million. Hank deposits the loan in his regular

checking account with First National. The deposits at First National rise by \$9 million and its reserves also rise by \$9 million, as [\[link\]](#) shows. First National must hold 10% of additional deposits as required reserves but is free to loan out the rest

First National Balance Sheet

Assets		Liabilities + Net Worth	
Reserves	+ \$9 million	Deposits	+ \$9 million

Making loans that are deposited into a demand deposit account increases the M1 money supply. Remember the definition of M1 includes checkable (demand) deposits, which can be easily used as a medium of exchange to buy goods and services. Notice that the money supply is now \$19 million: \$10 million in deposits in Singleton bank and \$9 million in deposits at First National. Obviously these deposits will be drawn down as Hank's Auto Supply writes checks to pay its bills. But the bigger picture is that a bank must hold enough money in reserves to meet its liabilities; the rest the bank loans out. In this example so far, bank lending has expanded the money supply by \$9 million.

Now, First National must hold only 10% as required reserves (\$900,000) but can lend out the other 90% (\$8.1 million) in a loan to Jack's Chevy Dealership as shown in [\[link\]](#).

First National Balance Sheet

Assets		Liabilities + Net Worth	
Reserves	\$900,000	Deposits	+ \$9 million
Loans	\$8.1 million		

If Jack's deposits the loan in its checking account at Second National, the money supply just increased by an additional \$8.1 million, as [\[link\]](#) shows.

Second National Bank's Balance Sheet

Assets		Liabilities + Net Worth	
Reserves	+ \$8.1 million	Deposits	+ \$8.1 million

How is this money creation possible? It is possible because there are multiple banks in the financial system, they are required to hold only a fraction of their deposits, and loans end up deposited in other banks, which increases deposits and, in essence, the money supply.

Note:

Watch this [video](#) to learn more about how banks create money.



The Money Multiplier and a Multi-Bank System

In a system with multiple banks, the initial excess reserve amount that Singleton Bank decided to lend to Hank's Auto Supply was deposited into Frist National Bank, which is free to loan out \$8.1 million. If all banks loan out their excess reserves, the money supply will expand. In a multi-bank system, the amount of money that the system can create is found by using the money multiplier. The money multiplier tells us by how many times a loan will be "multiplied" as it is spent in the economy and then re-deposited in other banks.

Fortunately, a formula exists for calculating the total of these many rounds of lending in a banking system. The **money multiplier formula** is:

Equation:

$$\frac{1}{\text{Reserve Requirement}}$$

The money multiplier is then multiplied by the change in excess reserves to determine the total amount of M1 money supply created in the banking system. See the Work it Out feature to walk through the multiplier calculation.

Note:

Using the Money Multiplier Formula

Using the money multiplier for the example in this text:

Step 1. In the case of Singleton Bank, for whom the reserve requirement is 10% (or 0.10), the money multiplier is 1 divided by .10, which is equal to 10.

Step 2. We have identified that the excess reserves are \$9 million, so, using the formula we can determine the total change in the M1 money supply:

Equation:

$$\begin{aligned}\text{Total Change in the M1 Money Supply} &= \frac{1}{\text{Reserve Requirement}} \times \text{Excess Requirement} \\ &= \frac{1}{0.10} \times \$9 \text{ million} \\ &= 10 \times \$9 \text{ million} \\ &= \$90 \text{ million}\end{aligned}$$

Step 3. Thus, we can say that, in this example, the total quantity of money generated in this economy after all rounds of lending are completed will be \$90 million.

Cautions about the Money Multiplier

The money multiplier will depend on the proportion of reserves that banks are required to hold by the Federal Reserve Bank. Additionally, a bank can also choose to hold extra reserves. Banks may decide to vary how much they hold in reserves for two reasons: macroeconomic conditions and government rules. When an economy is in recession, banks are likely to hold a higher proportion of reserves because they fear that loans are less likely to be repaid when the economy is slow. The Federal Reserve may also raise or lower the required reserves held by banks as a policy move to affect the quantity of money in an economy, as [Monetary Policy and Bank Regulation](#) will discuss.

The process of how banks create money shows how the quantity of money in an economy is closely linked to the quantity of lending or credit in the economy. Indeed, all of the money in the economy, except for the original reserves, is a result of bank loans that are re-deposited and loaned out, again, and again.

Finally, the money multiplier depends on people re-depositing the money that they receive in the banking system. If people instead store their cash in safe-deposit boxes or in shoeboxes hidden in their closets, then banks cannot recirculate the money in the form of loans. Indeed, central banks have an incentive to assure that bank deposits are safe because if people worry that they may lose their bank deposits, they may start holding more money in cash, instead of depositing it in banks, and the quantity of loans in an economy will decline. Low-income countries have what economists sometimes refer to as “mattress savings,” or money that people are hiding in their homes because they do not trust banks. When mattress savings in an economy are substantial, banks cannot lend out those funds and the money multiplier cannot operate as effectively. The overall quantity of money and loans in such an economy will decline.

Note:

Watch a [video](#) of Jem Bendell discussing “The Money Myth.”



Money and Banks—Benefits and Dangers

Money and banks are marvelous social inventions that help a modern economy to function. Compared with the alternative of barter, money makes market exchanges vastly easier in goods, labor, and financial markets. Banking makes money still more effective in facilitating exchanges in goods and labor markets. Moreover, the process of banks making loans in financial capital markets is intimately tied to the creation of money.

But the extraordinary economic gains that are possible through money and banking also suggest some possible corresponding dangers. If banks are not working well, it sets off a decline in convenience and safety of transactions throughout the economy. If the banks are under financial stress, because of a widespread decline in the value of their assets, loans may become far less available, which can deal a crushing blow to sectors of the economy that depend on borrowed money like business investment, home construction, and car manufacturing. The Great Recession of 2008–2009 illustrated this pattern.

Note:

The Many Disguises of Money: From Cowries to Bit Coins

The global economy has come a long way since it started using cowrie shells as currency. We have moved away from commodity and commodity-backed paper money to fiat currency. As technology and global integration increases, the need for paper currency is diminishing, too. Every day, we witness the increased use of debit and credit cards.

The latest creation and perhaps one of the purest forms of fiat money is the Bitcoin.

Bitcoins are a digital currency that allows users to buy goods and services online.

Products and services such as videos and books may be purchased using Bitcoins. It is not backed by any commodity nor has it been decreed by any government as legal tender, yet it used as a medium of exchange and its value (online at least) can be stored. It is also unregulated by any central bank, but is created online through people solving very complicated mathematics problems and getting paid afterward. Bitcoin.org is an information source if you are curious. Bitcoins are a relatively new type of money. At present, because it is not sanctioned as a legal currency by any country nor regulated by any central bank, it lends itself for use in illegal trading activities as well as legal ones. As technology increases and the need to reduce transactions costs associated with using traditional forms of money increases, Bitcoins or some sort of digital currency may replace our dollar bill, just as the cowrie shell was replaced.

Key Concepts and Summary

The money multiplier is defined as the quantity of money that the banking system can generate from each \$1 of bank reserves. The formula for calculating the multiplier is $1/\text{reserve ratio}$, where the reserve ratio is the fraction of deposits that the bank wishes to hold as reserves. The quantity of money in an economy and the quantity of credit for loans

are inextricably intertwined. Much of the money in an economy is created by the network of banks making loans, people making deposits, and banks making more loans.

Given the macroeconomic dangers of a malfunctioning banking system, [Monetary Policy and Bank Regulation](#) will discuss government policies for controlling the money supply and for keeping the banking system safe.

Self-Check Questions

Exercise:

Problem:

Imagine that you are in the position of buying loans in the secondary market (that is, buying the right to collect the payments on loans made by banks) for a bank or other financial services company. Explain why you would be willing to pay more or less for a given loan if:

- a. The borrower has been late on a number of loan payments
- b. Interest rates in the economy as a whole have risen since the loan was made
- c. The borrower is a firm that has just declared a high level of profits
- d. Interest rates in the economy as a whole have fallen since the loan was made

Solution:

- a. A borrower who has been late on a number of loan payments looks perhaps less likely to repay the loan, or to repay it on time, and so you would want to pay less for that loan.
- b. If interest rates generally have risen, then this loan made at a time of relatively lower interest rates looks less attractive, and you would pay less for it.
- c. If the borrower is a firm with a record of high profits, then it is likely to be able to repay the loan, and you would be willing to pay more for the loan.
- d. If interest rates in the economy have fallen, then the loan is worth more.

Review Questions

Exercise:

Problem: How do banks create money?

Exercise:

Problem: What is the formula for the money multiplier?

Critical Thinking Questions

Exercise:

Problem: Should banks have to hold 100% of their deposits? Why or why not?

Exercise:

Problem:

Explain what will happen to the money multiplier process if there is an increase in the reserve requirement?

Exercise:

Problem:

What do you think the Federal Reserve Bank did to the reserve requirement during the Great Recession of 2008–2009?

Problems

Exercise:

Problem:

Humongous Bank is the only bank in the economy. The people in this economy have \$20 million in money, and they deposit all their money in Humongous Bank.

- Humongous Bank decides on a policy of holding 100% reserves. Draw a T-account for the bank.
- Humongous Bank is required to hold 5% of its existing \$20 million as reserves, and to loan out the rest. Draw a T-account for the bank after this first round of loans has been made.
- Assume that Humongous bank is part of a multibank system. How much will money supply increase with that original loan of \$19 million?

References

Bitcoin. 2013. www.bitcoin.org.

National Public Radio. *Lawmakers and Regulators Take Closer Look at Bitcoin*. November 19, 2013. <http://thedianerehmsshow.org/shows/2013-11-19/lawmakers-and-regulators-take->

closer-look-bitcoin.

Glossary

money multiplier formula

total money in the economy divided by the original quantity of money, or change in
the total money in the economy divided by a change in the original quantity of money

Introduction to Monetary Policy and Bank Regulation

class="introduction"

Marriner S. Eccles Federal Reserve Headquarters, Washington D.C.

Some of the most
influential
decisions
regarding
monetary policy in
the United States
are made behind
these doors.

(Credit:
modification of
work by
“squirrel83”/Flick
r Creative
Commons)



Note:**The Problem of the Zero Percent Interest Rate Lower Bound**

Most economists believe that monetary policy (the manipulation of interest rates and credit conditions by a nation's central bank) has a powerful influence on a nation's economy. Monetary policy works when the central bank reduces interest rates and makes credit more available. As a result, business investment and other types of spending increase, causing GDP and employment to grow.

But what if the interest rates banks pay are close to zero already? They cannot be made negative, can they? That would mean that lenders pay borrowers for the privilege of taking their money. Yet, this was the situation the U.S. Federal Reserve found itself in at the end of the 2008–2009 recession. The federal funds rate, which is the interest rate for banks that the Federal Reserve targets with its monetary policy, was slightly above 5% in 2007. By 2009, it had fallen to 0.16%.

The Federal Reserve's situation was further complicated because fiscal policy, the other major tool for managing the economy, was constrained by fears that the federal budget deficit and the public debt were already too high. What were the Federal Reserve's options? How could monetary policy be used to stimulate the economy? The answer, as we will see in this chapter, was to change the rules of the game.

Note:**Introduction to Monetary Policy and Bank Regulation**

In this chapter, you will learn about:

- The Federal Reserve Banking System and Central Banks
- Bank Regulation
- How a Central Bank Executes Monetary Policy
- Monetary Policy and Economic Outcomes
- Pitfalls for Monetary Policy

Money, loans, and banks are all tied together. Money is deposited in bank accounts, which is then loaned to businesses, individuals, and other banks. When the interlocking system of money, loans, and banks works well, economic transactions are made smoothly in goods and labor markets and savers are connected with borrowers. If the money and banking system does not operate smoothly, the economy can either fall into recession or suffer prolonged inflation.

The government of every country has public policies that support the system of money, loans, and banking. But these policies do not always work perfectly. This chapter discusses how monetary policy works and what may prevent it from working perfectly.

The Federal Reserve Banking System and Central Banks

By the end of this section, you will be able to:

- Explain the structure and organization of the U.S. Federal Reserve
- Discuss how central banks impact monetary policy, promote financial stability, and provide banking services

In making decisions about the money supply, a central bank decides whether to raise or lower interest rates and, in this way, to influence macroeconomic policy, whose goal is low unemployment and low inflation. The central bank is also responsible for regulating all or part of the nation's banking system to protect bank depositors and insure the health of the bank's balance sheet.

The organization responsible for conducting monetary policy and ensuring that a nation's financial system operates smoothly is called the **central bank**. Most nations have central banks or currency boards. Some prominent central banks around the world include the European Central Bank, the Bank of Japan, and the Bank of England. In the United States, the central bank is called the Federal Reserve—often abbreviated as just “the Fed.” This section explains the organization of the U.S. Federal Reserve and identifies the major responsibilities of a central bank.

Structure/Organization of the Federal Reserve

Unlike most central banks, the Federal Reserve is semi-decentralized, mixing government appointees with representation from private-sector banks. At the national level, it is run by a Board of Governors, consisting of seven members appointed by the President of the United States and confirmed by the Senate. Appointments are for 14-year terms and they are arranged so that one term expires January 31 of every even-numbered year. The purpose of the long and staggered terms is to insulate the Board of Governors as much as possible from political pressure so that policy decisions can be made based only on their economic merits. Additionally, except when filling an unfinished term, each member only serves one term, further insulating decision-making from politics. Policy decisions of the

Fed do not require congressional approval, and the President cannot ask for the resignation of a Federal Reserve Governor as the President can with cabinet positions.

One member of the Board of Governors is designated as the Chair. For example, from 1987 until early 2006, the Chair was Alan Greenspan. From 2006 until 2014, Ben Bernanke held the post. The current Chair, Janet Yellen, has made many headlines already. Why? See the following Clear It Up feature to find out.

Note:

Who has the most immediate economic power in the world?
Chair of the Federal Reserve Board



Janet L. Yellen is the first woman to hold the position of Chair of the Federal Reserve Board of Governors.
(Credit: Board of Governors of the Federal Reserve System)

What individual can make financial market crash or soar just by making a public statement? It is not Bill Gates or Warren Buffett. It is not even the President of the United States. The answer is the Chair of the Federal Reserve Board of Governors. In early 2014, Janet L. Yellen, shown in [\[link\]](#) became the first woman to hold this post. Yellen has been described in the media as “perhaps the most qualified Fed chair in history.” With a Ph.D. in economics from Yale University, Yellen has taught macroeconomics at Harvard, the London School of Economics, and most recently at the University of California at Berkeley. From 2004–2010, Yellen was President of the Federal Reserve Bank of San Francisco. Not an ivory tower economist, Yellen became one of the few economists who warned about a possible bubble in the housing market, more than two years before the financial crisis occurred. Yellen served on the Board of Governors of the Federal Reserve twice, most recently as Vice Chair. She also spent two years as Chair of the President’s Council of Economic Advisors. If experience and credentials mean anything, Yellen is likely to be an effective Fed chair.

The Fed Chair is first among equals on the Board of Governors. While he or she has only one vote, the Chair controls the agenda, and is the public voice of the Fed, so he or she has more power and influence than one might expect.

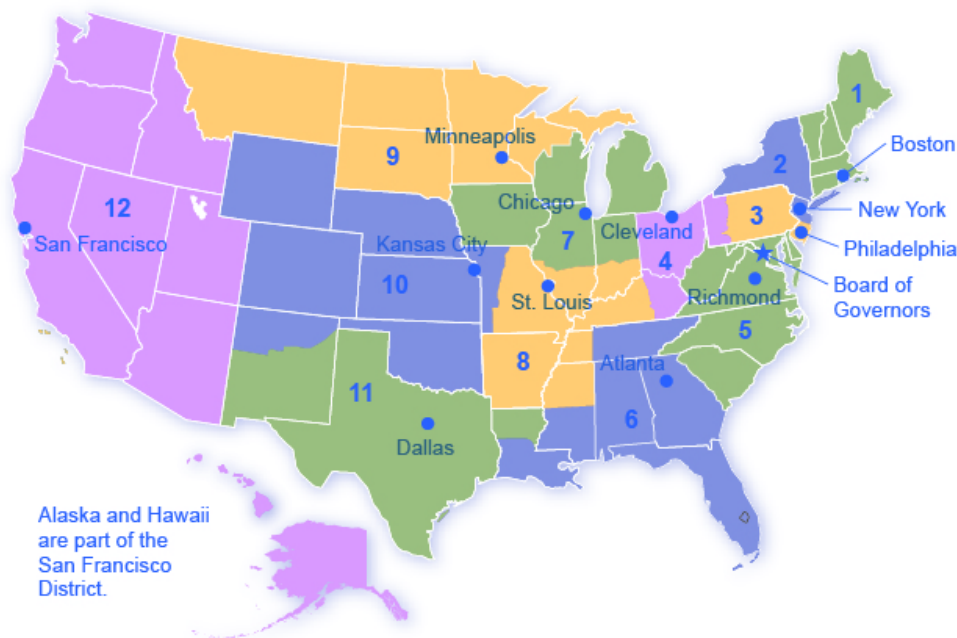
Note:

Visit this [website](#) to see who the current members of the Federal Reserve Board of Governors are. You can follow the links provided for each board member to learn more about their backgrounds, experiences, and when their terms on the board will end.



The Federal Reserve is more than the Board of Governors. The Fed also includes 12 regional Federal Reserve banks, each of which is responsible for supporting the commercial banks and economy generally in its district. The Federal Reserve districts and the cities where their regional headquarters are located are shown in [\[link\]](#). The commercial banks in each district elect a Board of Directors for each regional Federal Reserve bank, and that board chooses a president for each regional Federal Reserve district. Thus, the Federal Reserve System includes both federally and private-sector appointed leaders.

The Twelve Federal Reserve Districts



There are twelve regional Federal Reserve banks, each with its district.

What Does a Central Bank Do?

The Federal Reserve, like most central banks, is designed to perform three important functions:

1. To conduct monetary policy
2. To promote stability of the financial system
3. To provide banking services to commercial banks and other depository institutions, and to provide banking services to the federal government.

The first two functions are sufficiently important that we will discuss them in their own modules; the third function we will discuss here.

The Federal Reserve provides many of the same services to banks as banks provide to their customers. For example, all commercial banks have an account at the Fed where they deposit reserves. Similarly, banks can obtain loans from the Fed through the “discount window” facility, which will be discussed in more detail later. The Fed is also responsible for check processing. When you write a check, for example, to buy groceries, the grocery store deposits the check in its bank account. Then, the physical check (or an image of that actual check) is returned to your bank, after which funds are transferred from your bank account to the account of the grocery store. The Fed is responsible for each of these actions.

On a more mundane level, the Federal Reserve ensures that enough currency and coins are circulating through the financial system to meet public demands. For example, each year the Fed increases the amount of currency available in banks around the Christmas shopping season and reduces it again in January.

Finally, the Fed is responsible for assuring that banks are in compliance with a wide variety of consumer protection laws. For example, banks are forbidden from discriminating on the basis of age, race, sex, or marital status. Banks are also required to disclose publicly information about the loans they make for buying houses and how those loans are distributed geographically, as well as by sex and race of the loan applicants.

Key Concepts and Summary

The most prominent task of a central bank is to conduct monetary policy, which involves changes to interest rates and credit conditions, affecting the amount of borrowing and spending in an economy. Some prominent central banks around the world include the U.S. Federal Reserve, the European Central Bank, the Bank of Japan, and the Bank of England.

Self-Check Question

Exercise:

Problem:

Why is it important for the members of the Board of Governors of the Federal Reserve to have longer terms in office than elected officials, like the President?

Solution:

Longer terms insulate the Board from political forces. Since the presidency can potentially change every four years, the Federal Reserve's independence prevents drastic swings in monetary policy with every new administration and allows policy decisions to be made only on economic grounds.

Review Questions

Exercise:

Problem:

How is a central bank different from a typical commercial bank?

Exercise:

Problem:

List the three traditional tools that a central bank has for controlling the money supply.

Critical Thinking Questions**Exercise:****Problem:**

Why do presidents typically reappoint Chairs of the Federal Reserve Board even when they were originally appointed by a president of a different political party?

Exercise:**Problem:**

In what ways might monetary policy be superior to fiscal policy? In what ways might it be inferior?

References

Matthews, Dylan. "Nine amazing facts about Janet Yellen, our next Fed chair." Wonkblog. *The Washington Post*. Posted October 09, 2013.
<http://www.washingtonpost.com/blogs/wonkblog/wp/2013/10/09/nine-amazing-facts-about-janet-yellen-our-next-fed-chair/>.

Appelbaum, Binyamin. "Divining the Regulatory Goals of Fed Rivals." *The New York Times*. Posted August 14, 2013.
<http://www.nytimes.com/2013/08/14/business/economy/careers-of-2-fed-contenders-reveal-little-on-regulatory-approach.html?pagewanted=3>.

Glossary

central bank

institution which conducts a nation's monetary policy and regulates its banking system

Bank Regulation

By the end of this section, you will be able to:

- Discuss the relationship between bank regulation and monetary policy
- Explain bank supervision
- Explain how deposit insurance and lender of last resort are two strategies to protect against bank runs

A safe and stable national financial system is a critical concern of the Federal Reserve. The goal is not only to protect individuals' savings, but to protect the integrity of the financial system itself. This esoteric task is usually behind the scenes, but came into view during the 2008–2009 financial crisis, when for a brief period of time, critical parts of the financial system failed and firms became unable to obtain financing for ordinary parts of their business. Imagine if suddenly you were unable to access the money in your bank accounts because your checks were not accepted for payment and your debit cards were declined. This gives an idea of what a failure of the payments/financial system is like.

Bank regulation is intended to maintain the solvency of banks by avoiding excessive risk. Regulation falls into a number of categories, including reserve requirements, capital requirements, and restrictions on the types of investments banks may make. In [Money and Banking](#), we learned that banks are required to hold a minimum percentage of their deposits on hand as reserves. “On hand” is a bit of a misnomer because, while a portion of bank reserves are held as cash in the bank, the majority are held in the bank's account at the Federal Reserve, and their purpose is to cover desired withdrawals by depositors. Another part of bank regulation is restrictions on the types of investments banks are allowed to make. Banks are allowed to make loans to businesses, individuals, and other banks. They are allowed to purchase U.S. Treasury securities but, to protect depositors, they are not permitted to invest in the stock market or other assets that are perceived as too risky.

Bank capital is the difference between a bank's assets and its liabilities. In other words, it is a bank's net worth. A bank must have positive net worth; otherwise it is insolvent or bankrupt, meaning it would not have enough

assets to pay back its liabilities. Regulation requires that banks maintain a minimum net worth, usually expressed as a percent of their assets, to protect their depositors and other creditors.

Note:

Visit this [website](#) to read the brief article, “Stop Confusing Monetary Policy and Bank Regulation.”



Bank Supervision

Several government agencies monitor the balance sheets of banks to make sure they have positive net worth and are not taking too high a level of risk. Within the U.S. Department of the Treasury, the Office of the Comptroller of the Currency has a national staff of bank examiners who conduct on-site reviews of the 1,500 or so of the largest national banks. The bank examiners also review any foreign banks that have branches in the United States. The Office of the Comptroller of the Currency also monitors and regulates about 800 savings and loan institutions.

The National Credit Union Administration (NCUA) supervises credit unions, which are nonprofit banks owned and run by their members. There are over 6,000 credit unions in the U.S. economy, though the typical credit union is small compared to most banks.

The Federal Reserve also has some responsibility for supervising financial institutions. For example, conglomerate firms that own banks and other

businesses are called “bank holding companies.” While other regulators like the Office of the Comptroller of the Currency supervises the banks, the Federal Reserve supervises the holding companies.

When the supervision of banks (and bank-like institutions such as savings and loans and credit unions) works well, most banks will remain financially healthy most of the time. If the bank supervisors find that a bank has low or negative net worth, or is making too high a proportion of risky loans, they can require that the bank change its behavior—or, in extreme cases, even force the bank to be closed or sold to a financially healthy bank.

Bank supervision can run into both practical and political questions. The practical question is that measuring the value of a bank’s assets is not always straightforward. As discussed in [Money and Banking](#), a bank’s assets are its loans, and the value of these assets depends on estimates about the risk that these loans will not be repaid. These issues can become even more complex when a bank makes loans to banks or firms in other countries, or arranges financial deals that are much more complex than a basic loan.

The political question arises because the decision by a bank supervisor to require a bank to close or to change its financial investments is often controversial, and the bank supervisor often comes under political pressure from the owners of the bank and the local politicians to keep quiet and back off.

For example, many observers have pointed out that Japan’s banks were in deep financial trouble through most of the 1990s; however, nothing substantial had been done about it by the early 2000s. A similar unwillingness to confront problems with struggling banks is visible across the rest of the world, in East Asia, Latin America, Eastern Europe, Russia, and elsewhere.

In the United States, laws were passed in the 1990s requiring that bank supervisors make their findings open and public, and that they act as soon as a problem is identified. However, as many U.S. banks were staggered by the recession of 2008–2009, critics of the bank regulators asked pointed

questions about why the regulators had not foreseen the financial shakiness of the banks earlier, before such large losses had a chance to accumulate.

Bank Runs

Back in the nineteenth century and during the first few decades of the twentieth century (around and during the Great Depression), putting your money in a bank could be nerve-wracking. Imagine that the net worth of your bank became negative, so that the bank's assets were not enough to cover its liabilities. In this situation, whoever withdrew their deposits first received all of their money, and those who did not rush to the bank quickly enough, lost their money. Depositors racing to the bank to withdraw their deposits, as shown in [\[link\]](#) is called a **bank run**. In the movie *It's a Wonderful Life*, the bank manager, played by Jimmy Stewart, faces a mob of worried bank depositors who want to withdraw their money, but manages to allay their fears by allowing some of them to withdraw a portion of their deposits—using the money from his own pocket that was supposed to pay for his honeymoon.

A Run on the Bank



Bank runs during the Great Depression
only served to worsen the economic

situation. (Credit: National Archives and
Records Administration)

The risk of bank runs created instability in the banking system. Even a rumor that a bank might experience negative net worth could trigger a bank run and, in a bank run, even healthy banks could be destroyed. Because a bank loans out most of the money it receives, and because it keeps only limited reserves on hand, a bank run of any size would quickly drain any of the bank's available cash. When the bank had no cash remaining, it only intensified the fears of remaining depositors that they could lose their money. Moreover, a bank run at one bank often triggered a chain reaction of runs on other banks. In the late nineteenth and early twentieth century, bank runs were typically not the original cause of a recession—but they could make a recession much worse.

Deposit Insurance

To protect against bank runs, Congress has put two strategies into place: **deposit insurance** and the lender of last resort. Deposit insurance is an insurance system that makes sure depositors in a bank do not lose their money, even if the bank goes bankrupt. About 70 countries around the world, including all of the major economies, have deposit insurance programs. In the United States, the Federal Deposit Insurance Corporation (FDIC) is responsible for deposit insurance. Banks pay an insurance premium to the FDIC. The insurance premium is based on the bank's level of deposits, and then adjusted according to the riskiness of a bank's financial situation. In 2009, for example, a fairly safe bank with a high net worth might have paid 10–20 cents in insurance premiums for every \$100 in bank deposits, while a risky bank with very low net worth might have paid 50–60 cents for every \$100 in bank deposits.

Bank examiners from the FDIC evaluate the balance sheets of banks, looking at the value of assets and liabilities, to determine the level of riskiness. The FDIC provides deposit insurance for about 6,509 banks (as of the end of 2014). Even if a bank fails, the government guarantees that depositors will receive up to \$250,000 of their money in each account,

which is enough for almost all individuals, although not sufficient for many businesses. Since the United States enacted deposit insurance in the 1930s, no one has lost any of their insured deposits. Bank runs no longer happen at insured banks.

Lender of Last Resort

The problem with bank runs is not that insolvent banks will fail; they are, after all, bankrupt and need to be shut down. The problem is that bank runs can cause solvent banks to fail and spread to the rest of the financial system. To prevent this, the Fed stands ready to lend to banks and other financial institutions when they cannot obtain funds from anywhere else. This is known as the **lender of last resort** role. For banks, the central bank acting as a lender of last resort helps to reinforce the effect of deposit insurance and to reassure bank customers that they will not lose their money.

The lender of last resort task can come up in other financial crises, as well. During the panic of the stock market crash in 1987, when the value of U.S. stocks fell by 25% in a single day, the Federal Reserve made a number of short-term emergency loans so that the financial system could keep functioning. During the recession of 2008–2009, the “quantitative easing” policies (discussed below) of the Federal Reserve can be interpreted as a willingness to make short-term credit available as needed in a time when the banking and financial system was under stress.

Key Concepts and Summary

A bank run occurs when there are rumors (possibly true, possibly false) that a bank is at financial risk of having negative net worth. As a result, depositors rush to the bank to withdraw their money and put it someplace safer. Even false rumors, if they cause a bank run, can force a healthy bank to lose its deposits and be forced to close. Deposit insurance guarantees bank depositors that, even if the bank has negative net worth, their deposits will be protected. In the United States, the Federal Deposit Insurance Corporation (FDIC) collects deposit insurance premiums from banks and guarantees bank deposits up to \$250,000. Bank supervision involves inspecting the balance sheets of banks to make sure that they have positive

net worth and that their assets are not too risky. In the United States, the Office of the Comptroller of the Currency (OCC) is responsible for supervising banks and inspecting savings and loans and the National Credit Union Administration (NCUA) is responsible for inspecting credit unions. The FDIC and the Federal Reserve also play a role in bank supervision.

When a central bank acts as a lender of last resort, it makes short-term loans available in situations of severe financial panic or stress. The failure of a single bank can be treated like any other business failure. Yet if many banks fail, it can reduce aggregate demand in a way that can bring on or deepen a recession. The combination of deposit insurance, bank supervision, and lender of last resort policies help to prevent weaknesses in the banking system from causing recessions.

Self-Check Questions

Exercise:

Problem:

Given the danger of bank runs, why do banks not keep the majority of deposits on hand to meet the demands of depositors?

Solution:

Banks make their money from issuing loans and charging interest. The more money that is stored in the bank's vault, the less is available for lending and the less money the bank stands to make.

Exercise:

Problem:

Bank runs are often described as “self-fulfilling prophecies.” Why is this phrase appropriate to bank runs?

Solution:

The fear and uncertainty created by the suggestion that a bank might fail can lead depositors to withdraw their money. If many depositors do this at the same time, the bank may not be able to meet their demands and will, indeed, fail.

Review Questions

Exercise:

Problem:

How is bank regulation linked to the conduct of monetary policy?

Exercise:

Problem: What is a bank run?

Exercise:

Problem:

In a program of deposit insurance as it is operated in the United States, what is being insured and who pays the insurance premiums?

Exercise:

Problem:

In government programs of bank supervision, what is being supervised?

Exercise:

Problem: What is the lender of last resort?

Exercise:

Problem:

Name and briefly describe the responsibilities of each of the following agencies: FDIC, NCUA, and OCC.

Critical Thinking Question

Exercise:

Problem:

The term “moral hazard” describes increases in risky behavior resulting from efforts to make that behavior safer. How does the concept of moral hazard apply to deposit insurance and other bank regulations?

References

U.S. Department of the Treasury. “Office of the Comptroller of the Currency.” Accessed November 2013. <http://www.occ.gov/>.

National Credit Union Administration. “About NCUA.” Accessed November 2013. <http://www.ncua.gov/about/Pages/default.aspx>.

Glossary

bank run

when depositors race to the bank to withdraw their deposits for fear that otherwise they would be lost

deposit insurance

an insurance system that makes sure depositors in a bank do not lose their money, even if the bank goes bankrupt

lender of last resort

an institution that provides short-term emergency loans in conditions of financial crisis

How a Central Bank Executes Monetary Policy

By the end of this section, you will be able to:

- Explain the reason for open market operations
- Evaluate reserve requirements and discount rates
- Interpret and show bank activity through balance sheets

The most important function of the Federal Reserve is to conduct the nation's monetary policy. Article I, Section 8 of the U.S. Constitution gives Congress the power “to coin money” and “to regulate the value thereof.” As part of the 1913 legislation that created the Federal Reserve, Congress delegated these powers to the Fed. Monetary policy involves managing interest rates and credit conditions, which influences the level of economic activity, as described in more detail below.

A central bank has three traditional tools to implement monetary policy in the economy:

- Open market operations
- Changing reserve requirements
- Changing the discount rate

In discussing how these three tools work, it is useful to think of the central bank as a “bank for banks”—that is, each private-sector bank has its own account at the central bank. We will discuss each of these monetary policy tools in the sections below.

Open Market Operations

The most commonly used tool of monetary policy in the U.S. is **open market operations**. Open market operations take place when the central bank sells or buys U.S. Treasury bonds in order to influence the quantity of bank reserves and the level of interest rates. The specific interest rate targeted in open market operations is the federal funds rate. The name is a bit of a misnomer since the federal funds rate is the interest rate charged by commercial banks making overnight loans to other banks. As such, it is a

very short term interest rate, but one that reflects credit conditions in financial markets very well.

The **Federal Open Market Committee (FOMC)** makes the decisions regarding these open market operations. The FOMC is made up of the seven members of the Federal Reserve's Board of Governors. It also includes five voting members who are drawn, on a rotating basis, from the regional Federal Reserve Banks. The New York district president is a permanent voting member of the FOMC and the other four spots are filled on a rotating, annual basis, from the other 11 districts. The FOMC typically meets every six weeks, but it can meet more frequently if necessary. The FOMC tries to act by consensus; however, the chairman of the Federal Reserve has traditionally played a very powerful role in defining and shaping that consensus. For the Federal Reserve, and for most central banks, open market operations have, over the last few decades, been the most commonly used tool of monetary policy.

Note:

Visit this [website](#) for the Federal Reserve to learn more about current monetary policy.



To understand how open market operations affect the money supply, consider the balance sheet of Happy Bank, displayed in [\[link\]](#). [\[link\]](#) (a) shows that Happy Bank starts with \$460 million in assets, divided among reserves, bonds and loans, and \$400 million in liabilities in the form of deposits, with a net worth of \$60 million. When the central bank purchases

\$20 million in bonds from Happy Bank, the bond holdings of Happy Bank fall by \$20 million and the bank's reserves rise by \$20 million, as shown in [\[link\]](#) (b). However, Happy Bank only wants to hold \$40 million in reserves (the quantity of reserves that it started with in [\[link\]](#)) (a), so the bank decides to loan out the extra \$20 million in reserves and its loans rise by \$20 million, as shown in [\[link\]](#) (c). The open market operation by the central bank causes Happy Bank to make loans instead of holding its assets in the form of government bonds, which expands the money supply. As the new loans are deposited in banks throughout the economy, these banks will, in turn, loan out some of the deposits they receive, triggering the money multiplier discussed in [Money and Banking](#).

Assets		Liabilities + Net Worth	
Reserves	40	Deposits	400
Bonds	120		
Loans	300	Net Worth	60

(a) The original balance sheet

Assets		Liabilities + Net Worth	
Reserves	$40 + 20 = 60$	Deposits	400
Bonds	$120 - 20 = 100$		
Loans	300	Net Worth	60

(b) The central bank buys bonds

Assets		Liabilities + Net Worth	
Reserves	$60 - 20 = 40$	Deposits	400
Bonds	100		
Loans	$300 + 20 = 320$	Net Worth	60

(c) The bank makes additional loans

Where did the Federal Reserve get the \$20 million that it used to purchase the bonds? A central bank has the power to create money. In practical terms, the Federal Reserve would write a check to Happy Bank, so that Happy Bank can have that money credited to its bank account at the Federal Reserve. In truth, the Federal Reserve created the money to purchase the bonds out of thin air—or with a few clicks on some computer keys.

Open market operations can also reduce the quantity of money and loans in an economy. [\[link\]](#) (a) shows the balance sheet of Happy Bank before the central bank sells bonds in the open market. When Happy Bank purchases \$30 million in bonds, Happy Bank sends \$30 million of its reserves to the central bank, but now holds an additional \$30 million in bonds, as shown in [\[link\]](#) (b). However, Happy Bank wants to hold \$40 million in reserves, as in [\[link\]](#) (a), so it will adjust down the quantity of its loans by \$30 million, to bring its reserves back to the desired level, as shown in [\[link\]](#) (c). In practical terms, a bank can easily reduce its quantity of loans. At any given time, a bank is receiving payments on loans that it made previously and also making new loans. If the bank just slows down or briefly halts making new loans, and instead adds those funds to its reserves, then its overall quantity of loans will decrease. A decrease in the quantity of loans also means fewer deposits in other banks, and other banks reducing their lending as well, as the money multiplier discussed in [Money and Banking](#) takes effect. And what about all those bonds? How do they affect the money supply? Read the following Clear It Up feature for the answer.

Assets		Liabilities + Net Worth	
Reserves	40	Deposits	400
Bonds	120		
Loans	300	Net Worth	60

(a) The original balance sheet

Assets		Liabilities + Net Worth	
Reserves	$40 - 30 = 10$	Deposits	400
Bonds	$120 + 30 = 150$		
Loans	300	Net Worth	60

(b) The central bank sells bonds to the bank

Assets		Liabilities + Net Worth	
Reserves	$10 + 30 = 40$	Deposits	400
Bonds	150		
Loans	$300 - 30 = 270$	Net Worth	60

(c) The bank makes fewer loans

Note:

Does selling or buying bonds increase the money supply?

Is it a sale of bonds by the central bank which increases bank reserves and lowers interest rates or is it a purchase of bonds by the central bank? The easy way to keep track of this is to treat the central bank as being *outside* the banking system. When a central bank buys bonds, money is flowing from the central bank to individual banks in the economy, increasing the supply of money in circulation. When a central bank sells bonds, then money from individual banks in the economy is flowing into the central bank—reducing the quantity of money in the economy.

Changing Reserve Requirements

A second method of conducting monetary policy is for the central bank to raise or lower the **reserve requirement**, which, as we noted earlier, is the percentage of each bank's deposits that it is legally required to hold either as cash in their vault or on deposit with the central bank. If banks are required to hold a greater amount in reserves, they have less money available to lend out. If banks are allowed to hold a smaller amount in reserves, they will have a greater amount of money available to lend out.

In early 2015, the Federal Reserve required banks to hold reserves equal to 0% of the first \$14.5 million in deposits, then to hold reserves equal to 3% of the deposits up to \$103.6 million, and 10% of any amount above \$103.6 million. Small changes in the reserve requirements are made almost every year. For example, the \$103.6 million dividing line is sometimes bumped up or down by a few million dollars. In practice, large changes in reserve requirements are rarely used to execute monetary policy. A sudden demand that all banks increase their reserves would be extremely disruptive and difficult to comply with, while loosening requirements too much would create a danger of banks being unable to meet the demand for withdrawals.

Changing the Discount Rate

The Federal Reserve was founded in the aftermath of the Financial Panic of 1907 when many banks failed as a result of bank runs. As mentioned earlier, since banks make profits by lending out their deposits, no bank, even those that are not bankrupt, can withstand a bank run. As a result of the Panic, the Federal Reserve was founded to be the “lender of last resort.” In the event of a bank run, sound banks, (banks that were not bankrupt) could borrow as much cash as they needed from the Fed’s discount “window” to quell the bank run. The interest rate banks pay for such loans is called the **discount rate**. (They are so named because loans are made against the bank’s outstanding loans “at a discount” of their face value.) Once depositors became convinced that the bank would be able to honor their withdrawals, they no longer had a reason to make a run on the bank. In short, the Federal Reserve was originally intended to provide credit passively, but in the years since its founding, the Fed has taken on a more active role with monetary policy.

So, the third traditional method for conducting monetary policy is to raise or lower the discount rate. If the central bank raises the discount rate, then commercial banks will reduce their borrowing of reserves from the Fed, and instead call in loans to replace those reserves. Since fewer loans are available, the money supply falls and market interest rates rise. If the central bank lowers the discount rate it charges to banks, the process works in reverse.

In recent decades, the Federal Reserve has made relatively few discount loans. Before a bank borrows from the Federal Reserve to fill out its required reserves, the bank is expected to first borrow from other available sources, like other banks. This is encouraged by Fed’s charging a higher discount rate, than the federal funds rate. Given that most banks borrow little at the discount rate, changing the discount rate up or down has little impact on their behavior. More importantly, the Fed has found from experience that open market operations are a more precise and powerful means of executing any desired monetary policy.

In the Federal Reserve Act, the phrase “...to afford means of rediscounting commercial paper” is contained in its long title. This tool was seen as the

main tool for monetary policy when the Fed was initially created. This illustrates how monetary policy has evolved and how it continues to do so.

Key Concepts and Summary

A central bank has three traditional tools to conduct monetary policy: open market operations, which involves buying and selling government bonds with banks; reserve requirements, which determine what level of reserves a bank is legally required to hold; and discount rates, which is the interest rate charged by the central bank on the loans that it gives to other commercial banks. The most commonly used tool is open market operations.

Self-Check Questions

Exercise:

Problem:

If the central bank sells \$500 in bonds to a bank that has issued \$10,000 in loans and is exactly meeting the reserve requirement of 10%, what will happen to the amount of loans and to the money supply in general?

Solution:

The bank has to hold \$1,000 in reserves, so when it buys the \$500 in bonds, it will have to reduce its loans by \$500 to make up the difference. The money supply decreases by the same amount.

Exercise:

Problem:

What would be the effect of increasing the reserve requirements of banks on the money supply?

Solution:

An increase in reserve requirements would reduce the supply of money, since more money would be held in banks rather than circulating in the economy.

Review Questions

Exercise:

Problem:

Explain how to use an open market operation to expand the money supply.

Exercise:

Problem:

Explain how to use the reserve requirement to expand the money supply.

Exercise:

Problem:

Explain how to use the discount rate to expand the money supply.

Critical Thinking Question

Exercise:

Problem:

Explain what would happen if banks were notified they had to increase their required reserves by one percentage point from, say, 9% to 10% of deposits. What would their options be to come up with the cash?

Problems

Exercise:

Problem:

Suppose the Fed conducts an open market purchase by buying \$10 million in Treasury bonds from Acme Bank. Sketch out the balance sheet changes that will occur as Acme converts the bond sale proceeds to new loans. The initial Acme bank balance sheet contains the following information: Assets – reserves 30, bonds 50, and loans 50; Liabilities – deposits 300 and equity 30.

Exercise:**Problem:**

Suppose the Fed conducts an open market sale by selling \$10 million in Treasury bonds to Acme Bank. Sketch out the balance sheet changes that will occur as Acme restores its required reserves (10% of deposits) by reducing its loans. The initial balance sheet for Acme Bank contains the following information: Assets – reserves 30, bonds 50, and loans 250; Liabilities – deposits 300 and equity 30.

References

Board of Governors of the Federal Reserve System. "Federal Open Market Committee." Accessed September 3, 2013.

<http://www.federalreserve.gov/monetarypolicy/fomc.htm>.

Board of Governors of the Federal Reserve System. "Reserve Requirements." Accessed November 5, 2013.

<http://www.federalreserve.gov/monetarypolicy/reservereq.htm>.

Cox, Jeff. 2014. "Fed Completes the Taper." Accessed March 31, 2015.

<http://www.cnn.com/id/102132961>.

Jahan, Sarwat. n.d. "Inflation Targeting: Holding the Line." International Monetary Fund. Accessed March 31, 2015.

<http://www.imf.org/external/pubs/ft/fandd/basics/target.htm>.

Glossary

discount rate

the interest rate charged by the central bank on the loans that it gives to other commercial banks

open market operations

the central bank selling or buying Treasury bonds to influence the quantity of money and the level of interest rates

reserve requirement

the percentage amount of its total deposits that a bank is legally obligated to either hold as cash in their vault or deposit with the central bank

Monetary Policy and Economic Outcomes

By the end of this section, you will be able to:

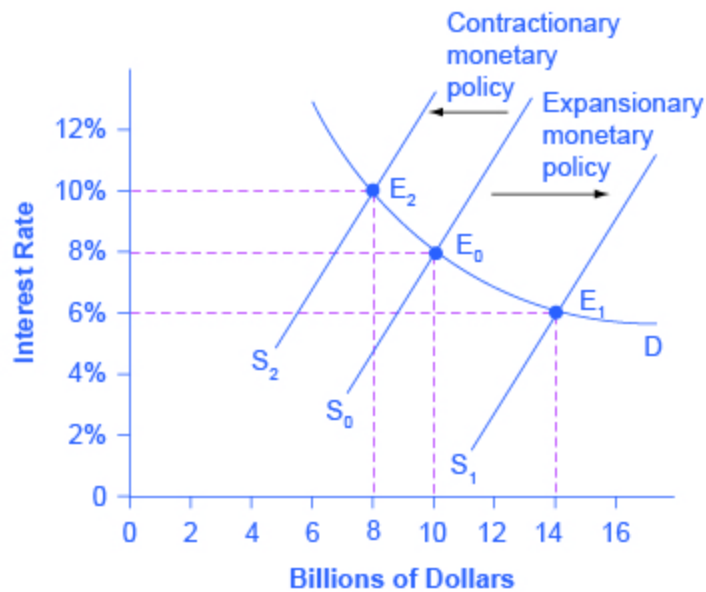
- Contrast expansionary monetary policy and contractionary monetary policy
- Explain how monetary policy impacts interest rates and aggregate demand
- Evaluate Federal Reserve decisions over the last forty years
- Explain the significance of quantitative easing (QE)

A monetary policy that lowers interest rates and stimulates borrowing is known as an **expansionary monetary policy** or **loose monetary policy**. Conversely, a monetary policy that raises interest rates and reduces borrowing in the economy is a **contractionary monetary policy** or **tight monetary policy**. This module will discuss how expansionary and contractionary monetary policies affect interest rates and aggregate demand, and how such policies will affect macroeconomic goals like unemployment and inflation. We will conclude with a look at the Fed's monetary policy practice in recent decades.

The Effect of Monetary Policy on Interest Rates

Consider the market for loanable bank funds, shown in [\[link\]](#). The original equilibrium (E_0) occurs at an interest rate of 8% and a quantity of funds loaned and borrowed of \$10 billion. An expansionary monetary policy will shift the supply of loanable funds to the right from the original supply curve (S_0) to S_1 , leading to an equilibrium (E_1) with a lower interest rate of 6% and a quantity of funds loaned of \$14 billion. Conversely, a contractionary monetary policy will shift the supply of loanable funds to the left from the original supply curve (S_0) to S_2 , leading to an equilibrium (E_2) with a higher interest rate of 10% and a quantity of funds loaned of \$8 billion.

Monetary Policy and Interest Rates



The original equilibrium occurs at E_0 . An expansionary monetary policy will shift the supply of loanable funds to the right from the original supply curve (S_0) to the new supply curve (S_1) and to a new equilibrium of E_1 , reducing the interest rate from 8% to 6%. A contractionary monetary policy will shift the supply of loanable funds to the left from the original supply curve (S_0) to the new supply (S_2), and raise the interest rate from 8% to 10%.

So how does a central bank “raise” interest rates? When describing the monetary policy actions taken by a central bank, it is common to hear that the central bank “raised interest rates” or “lowered interest rates.” We need to be clear about this: more precisely, through open market operations the central bank changes bank reserves in a way which affects the supply curve of loanable funds. As a result, interest rates change, as shown in [\[link\]](#). If they do not meet the Fed’s target, the Fed can supply more or less reserves until interest rates do.

Recall that the specific interest rate the Fed targets is the **federal funds rate**. The Federal Reserve has, since 1995, established its target federal funds rate in advance of any open market operations.

Of course, financial markets display a wide range of interest rates, representing borrowers with different risk premiums and loans that are to be repaid over different periods of time. In general, when the federal funds rate drops substantially, other interest rates drop, too, and when the federal funds rate rises, other interest rates rise. However, a fall or rise of one percentage point in the federal funds rate—which remember is for borrowing overnight—will typically have an effect of less than one percentage point on a 30-year loan to purchase a house or a three-year loan to purchase a car. Monetary policy can push the entire spectrum of interest rates higher or lower, but the specific interest rates are set by the forces of supply and demand in those specific markets for lending and borrowing.

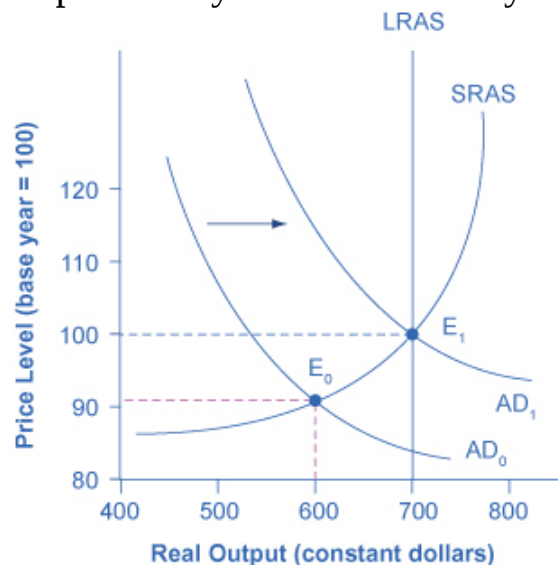
The Effect of Monetary Policy on Aggregate Demand

Monetary policy affects interest rates and the available quantity of loanable funds, which in turn affects several components of aggregate demand. Tight or contractionary monetary policy that leads to higher interest rates and a reduced quantity of loanable funds will reduce two components of aggregate demand. Business investment will decline because it is less attractive for firms to borrow money, and even firms that have money will notice that, with higher interest rates, it is relatively more attractive to put those funds in a financial investment than to make an investment in physical capital. In addition, higher interest rates will discourage consumer borrowing for big-ticket items like houses and cars. Conversely, loose or expansionary monetary policy that leads to lower interest rates and a higher quantity of loanable funds will tend to increase business investment and consumer borrowing for big-ticket items.

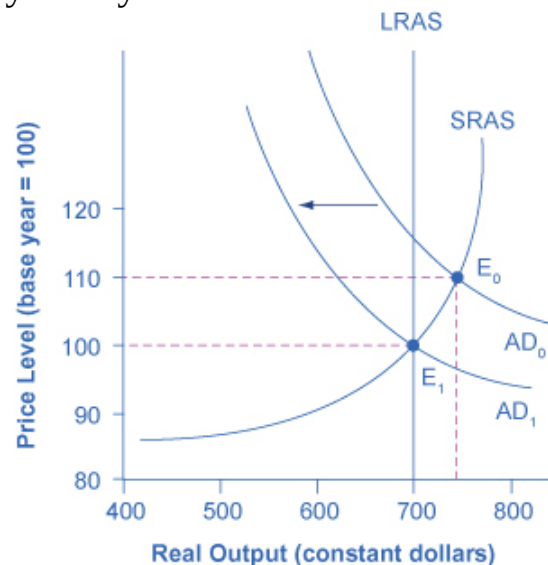
If the economy is suffering a recession and high unemployment, with output below potential GDP, expansionary monetary policy can help the economy return to potential GDP. [\[link\]](#) (a) illustrates this situation. This example uses a short-run upward-sloping Keynesian aggregate supply curve (SRAS). The original equilibrium during a recession of E_0 occurs at an output level

of 600. An expansionary monetary policy will reduce interest rates and stimulate investment and consumption spending, causing the original aggregate demand curve (AD_0) to shift right to AD_1 , so that the new equilibrium (E_1) occurs at the potential GDP level of 700.

Expansionary or Contractionary Monetary Policy



(a) Expansionary monetary policy



(b) Contractionary monetary policy

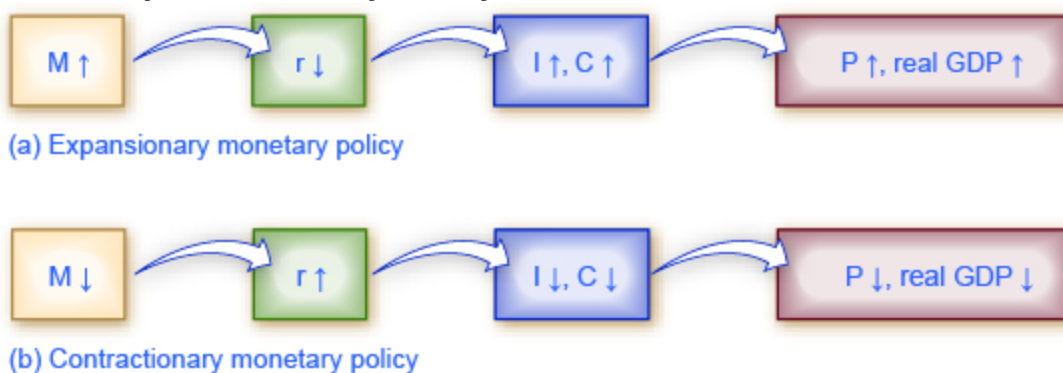
(a) The economy is originally in a recession with the equilibrium output and price level shown at E_0 . Expansionary monetary policy will reduce interest rates and shift aggregate demand to the right from AD_0 to AD_1 , leading to the new equilibrium (E_1) at the potential GDP level of output with a relatively small rise in the price level. (b) The economy is originally producing above the potential GDP level of output at the equilibrium E_0 and is experiencing pressures for an inflationary rise in the price level. Contractionary monetary policy will shift aggregate demand to the left from AD_0 to AD_1 , thus leading to a new equilibrium (E_1) at the potential GDP level of output.

Conversely, if an economy is producing at a quantity of output above its potential GDP, a contractionary monetary policy can reduce the inflationary pressures for a rising price level. In [\[link\]](#) (b), the original equilibrium (E_0) occurs at an output of 750, which is above potential GDP. A contractionary

monetary policy will raise interest rates, discourage borrowing for investment and consumption spending, and cause the original demand curve (AD_0) to shift left to AD_1 , so that the new equilibrium (E_1) occurs at the potential GDP level of 700.

These examples suggest that monetary policy should be **countercyclical**; that is, it should act to counterbalance the business cycles of economic downturns and upswings. Monetary policy should be loosened when a recession has caused unemployment to increase and tightened when inflation threatens. Of course, countercyclical policy does pose a danger of overreaction. If loose monetary policy seeking to end a recession goes too far, it may push aggregate demand so far to the right that it triggers inflation. If tight monetary policy seeking to reduce inflation goes too far, it may push aggregate demand so far to the left that a recession begins. [\[link\]](#) (a) summarizes the chain of effects that connect loose and tight monetary policy to changes in output and the price level.

The Pathways of Monetary Policy



(a) In expansionary monetary policy the central bank causes the supply of money and loanable funds to increase, which lowers the interest rate, stimulating additional borrowing for investment and consumption, and shifting aggregate demand right. The result is a higher price level and, at least in the short run, higher real GDP. (b)

In contractionary monetary policy, the central bank causes the supply of money and credit in the economy to decrease, which raises the interest rate, discouraging borrowing for investment and consumption, and shifting aggregate demand left. The result is a lower price level and, at least in the short run, lower real GDP.

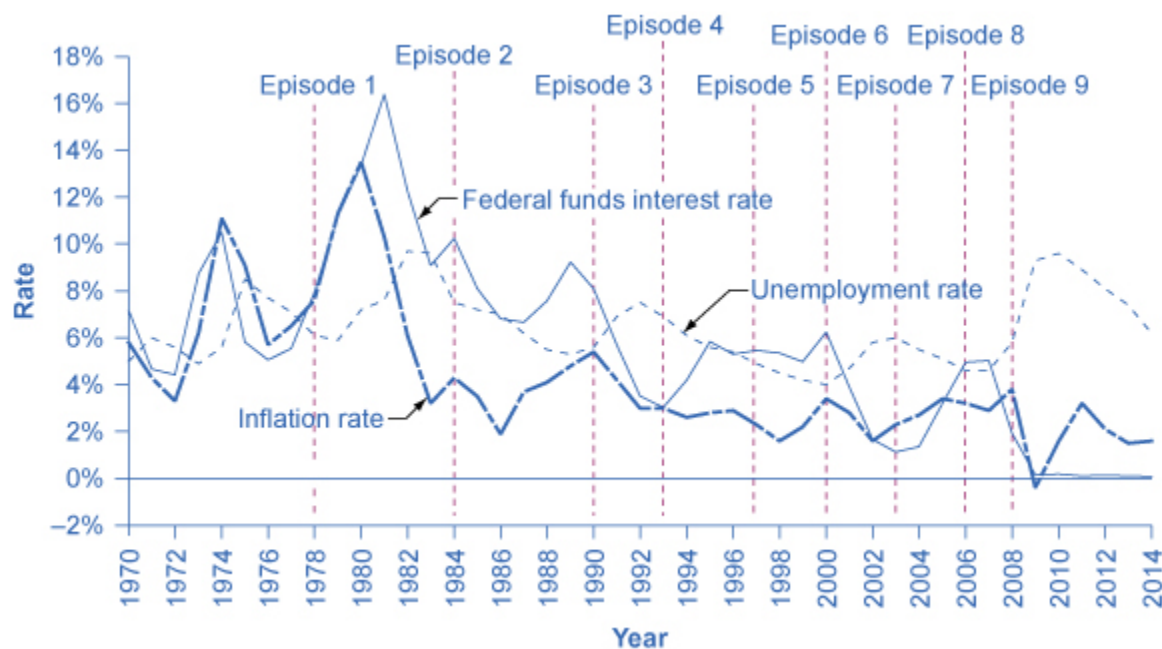
Federal Reserve Actions Over Last Four Decades

For the period from the mid-1970s up through the end of 2007, Federal Reserve monetary policy can largely be summed up by looking at how it targeted the federal funds interest rate using open market operations.

Of course, telling the story of the U.S. economy since 1975 in terms of Federal Reserve actions leaves out many other macroeconomic factors that were influencing unemployment, recession, economic growth, and inflation over this time. The nine episodes of Federal Reserve action outlined in the sections below also demonstrate that the central bank should be considered one of the leading actors influencing the macro economy. As noted earlier, the single person with the greatest power to influence the U.S. economy is probably the chairperson of the Federal Reserve.

[\[link\]](#) shows how the Federal Reserve has carried out monetary policy by targeting the federal funds interest rate in the last few decades. The graph shows the federal funds interest rate (remember, this interest rate is set through open market operations), the unemployment rate, and the inflation rate since 1975. Different episodes of monetary policy during this period are indicated in the figure.

Monetary Policy, Unemployment, and Inflation



Through the episodes shown here, the Federal Reserve typically reacted to higher inflation with a contractionary monetary policy and a higher interest rate, and reacted to higher unemployment with an expansionary monetary policy and a lower interest rate.

Episode 1

Consider Episode 1 in the late 1970s. The rate of inflation was very high, exceeding 10% in 1979 and 1980, so the Federal Reserve used tight monetary policy to raise interest rates, with the federal funds rate rising from 5.5% in 1977 to 16.4% in 1981. By 1983, inflation was down to 3.2%, but aggregate demand contracted sharply enough that back-to-back recessions occurred in 1980 and in 1981–1982, and the unemployment rate rose from 5.8% in 1979 to 9.7% in 1982.

Episode 2

In Episode 2, when the Federal Reserve was persuaded in the early 1980s that inflation was declining, the Fed began slashing interest rates to reduce unemployment. The federal funds interest rate fell from 16.4% in 1981 to

6.8% in 1986. By 1986 or so, inflation had fallen to about 2% and the unemployment rate had come down to 7%, and was still falling.

Episode 3

However, in Episode 3 in the late 1980s, inflation appeared to be creeping up again, rising from 2% in 1986 up toward 5% by 1989. In response, the Federal Reserve used contractionary monetary policy to raise the federal funds rates from 6.6% in 1987 to 9.2% in 1989. The tighter monetary policy stopped inflation, which fell from above 5% in 1990 to under 3% in 1992, but it also helped to cause the recession of 1990–1991, and the unemployment rate rose from 5.3% in 1989 to 7.5% by 1992.

Episode 4

In Episode 4, in the early 1990s, when the Federal Reserve was confident that inflation was back under control, it reduced interest rates, with the federal funds interest rate falling from 8.1% in 1990 to 3.5% in 1992. As the economy expanded, the unemployment rate declined from 7.5% in 1992 to less than 5% by 1997.

Episodes 5 and 6

In Episodes 5 and 6, the Federal Reserve perceived a risk of inflation and raised the federal funds rate from 3% to 5.8% from 1993 to 1995. Inflation did not rise, and the period of economic growth during the 1990s continued. Then in 1999 and 2000, the Fed was concerned that inflation seemed to be creeping up so it raised the federal funds interest rate from 4.6% in December 1998 to 6.5% in June 2000. By early 2001, inflation was declining again, but a recession occurred in 2001. Between 2000 and 2002, the unemployment rate rose from 4.0% to 5.8%.

Episodes 7 and 8

In Episodes 7 and 8, the Federal Reserve conducted a loose monetary policy and slashed the federal funds rate from 6.2% in 2000 to just 1.7% in 2002, and then again to 1% in 2003. They actually did this because of fear of Japan-style deflation; this persuaded them to lower the Fed funds further

than they otherwise would have. The recession ended, but, unemployment rates were slow to decline in the early 2000s. Finally, in 2004, the unemployment rate declined and the Federal Reserve began to raise the federal funds rate until it reached 5% by 2007.

Episode 9

In Episode 9, as the Great Recession took hold in 2008, the Federal Reserve was quick to slash interest rates, taking them down to 2% in 2008 and to nearly 0% in 2009. When the Fed had taken interest rates down to near-zero by December 2008, the economy was still deep in recession. Open market operations could not make the interest rate turn negative. The Federal Reserve had to think “outside the box.”

Quantitative Easing

The most powerful and commonly used of the three traditional tools of monetary policy—open market operations—works by expanding or contracting the money supply in a way that influences the interest rate. In late 2008, as the U.S. economy struggled with recession, the Federal Reserve had already reduced the interest rate to near-zero. With the recession still ongoing, the Fed decided to adopt an innovative and nontraditional policy known as **quantitative easing (QE)**. This is the purchase of long-term government and private mortgage-backed securities by central banks to make credit available so as to stimulate aggregate demand.

Quantitative easing differed from traditional monetary policy in several key ways. First, it involved the Fed purchasing long term Treasury bonds, rather than short term Treasury bills. In 2008, however, it was impossible to stimulate the economy any further by lowering short term rates because they were already as low as they could get. (Read the closing Bring it Home feature for more on this.) Therefore, Bernanke sought to lower long-term rates utilizing quantitative easing.

This leads to a second way QE is different from traditional monetary policy. Instead of purchasing Treasury securities, the Fed also began purchasing

private mortgage-backed securities, something it had never done before. During the financial crisis, which precipitated the recession, mortgage-backed securities were termed “toxic assets,” because when the housing market collapsed, no one knew what these securities were worth, which put the financial institutions which were holding those securities on very shaky ground. By offering to purchase mortgage-backed securities, the Fed was both pushing long term interest rates down and also removing possibly “toxic assets” from the balance sheets of private financial firms, which would strengthen the financial system.

Quantitative easing (QE) occurred in three episodes:

1. During QE₁, which began in November 2008, the Fed purchased \$600 billion in mortgage-backed securities from government enterprises Fannie Mae and Freddie Mac.
2. In November 2010, the Fed began QE₂, in which it purchased \$600 billion in U.S. Treasury bonds.
3. QE₃, began in September 2012 when the Fed commenced purchasing \$40 billion of additional mortgage-backed securities per month. This amount was increased in December 2012 to \$85 billion per month. The Fed stated that, when economic conditions permit, it will begin tapering (or reducing the monthly purchases). By October 2014, the Fed had announced the final \$15 billion purchase of bonds, ending Quantitative Easing.

The quantitative easing policies adopted by the Federal Reserve (and by other central banks around the world) are usually thought of as temporary emergency measures. If these steps are, indeed, to be temporary, then the Federal Reserve will need to stop making these additional loans and sell off the financial securities it has accumulated. The concern is that the process of quantitative easing may prove more difficult to reverse than it was to enact. The evidence suggests that QE₁ was somewhat successful, but that QE₂ and QE₃ have been less so.

Key Concepts and Summary

An expansionary (or loose) monetary policy raises the quantity of money and credit above what it otherwise would have been and reduces interest rates, boosting aggregate demand, and thus countering recession. A contractionary monetary policy, also called a tight monetary policy, reduces the quantity of money and credit below what it otherwise would have been and raises interest rates, seeking to hold down inflation. During the 2008–2009 recession, central banks around the world also used quantitative easing to expand the supply of credit.

Self-Check Questions

Exercise:

Problem:

Why does contractionary monetary policy cause interest rates to rise?

Solution:

Contractionary policy reduces the amount of loanable funds in the economy. As with all goods, greater scarcity leads a greater price, so the interest rate, or the price of borrowing money, rises.

Exercise:

Problem:

Why does expansionary monetary policy causes interest rates to drop?

Solution:

An increase in the amount of available loanable funds means that there are more people who want to lend. They, therefore, bid the price of borrowing (the interest rate) down.

Review Questions

Exercise:

Problem:

How do the expansionary and contractionary monetary policy affect the quantity of money?

Exercise:**Problem:**

How do tight and loose monetary policy affect interest rates?

Exercise:**Problem:**

How do expansionary, tight, contractionary, and loose monetary policy affect aggregate demand?

Exercise:**Problem:**

Which kind of monetary policy would you expect in response to high inflation: expansionary or contractionary? Why?

Exercise:**Problem:**

Explain how to use quantitative easing to stimulate aggregate demand.

Critical Thinking Question**Exercise:****Problem:**

A well-known economic model called the Phillips Curve (discussed in [The Keynesian Perspective](#) chapter) describes the short run tradeoff typically observed between inflation and unemployment. Based on the discussion of expansionary and contractionary monetary policy, explain why one of these variables usually falls when the other rises.

Glossary

contractionary monetary policy

a monetary policy that reduces the supply of money and loans

countercyclical

moving in the opposite direction of the business cycle of economic downturns and upswings

expansionary monetary policy

a monetary policy that increases the supply of money and the quantity of loans

federal funds rate

the interest rate at which one bank lends funds to another bank overnight

loose monetary policy

see expansionary monetary policy

quantitative easing (QE)

the purchase of long term government and private mortgage-backed securities by central banks to make credit available in hopes of stimulating aggregate demand

tight monetary policy

see contractionary monetary policy

Pitfalls for Monetary Policy

By the end of this section, you will be able to:

- Analyze whether monetary policy decisions should be made more democratically
- Calculate the velocity of money
- Evaluate the central bank's influence on inflation, unemployment, asset bubbles, and leverage cycles
- Calculate the effects of monetary stimulus

In the real world, effective monetary policy faces a number of significant hurdles. Monetary policy affects the economy only after a time lag that is typically long and of variable length. Remember, monetary policy involves a chain of events: the central bank must perceive a situation in the economy, hold a meeting, and make a decision to react by tightening or loosening monetary policy. The change in monetary policy must percolate through the banking system, changing the quantity of loans and affecting interest rates. When interest rates change, businesses must change their investment levels and consumers must change their borrowing patterns when purchasing homes or cars. Then it takes time for these changes to filter through the rest of the economy.

As a result of this chain of events, monetary policy has little effect in the immediate future; instead, its primary effects are felt perhaps one to three years in the future. The reality of long and variable time lags does not mean that a central bank should refuse to make decisions. It does mean that central banks should be humble about taking action, because of the risk that their actions can create as much or more economic instability as they resolve.

Excess Reserves

Banks are legally required to hold a minimum level of reserves, but no rule prohibits them from holding additional **excess reserves** above the legally mandated limit. For example, during a recession banks may be hesitant to lend, because they fear that when the economy is contracting, a high proportion of loan applicants become less likely to repay their loans.

When many banks are choosing to hold excess reserves, expansionary monetary policy may not work well. This may occur because the banks are concerned about a deteriorating economy, while the central bank is trying to expand the money supply. If the banks prefer to hold excess reserves above the legally required level, the central bank cannot force individual banks to make loans. Similarly, sensible businesses and consumers may be reluctant to borrow substantial amounts of money in a recession, because they recognize that firms' sales and employees' jobs are more insecure in a recession, and they do not want to face the need to make interest payments. The result is that during an especially deep recession, an expansionary monetary policy may have little effect on either the price level or the real GDP.

Japan experienced this situation in the 1990s and early 2000s. Japan's economy entered a period of very slow growth, dipping in and out of recession, in the early 1990s. By February 1999, the Bank of Japan had lowered the equivalent of its federal funds rate to 0%. It kept it there most of the time through 2003. Moreover, in the two years from March 2001 to March 2003, the Bank of Japan also expanded the money supply of the country by about 50%—an enormous increase. Even this highly expansionary monetary policy, however, had no substantial effect on stimulating aggregate demand. Japan's economy continued to experience extremely slow growth into the mid-2000s.

Note:

Should monetary policy decisions be made more democratically?

Should monetary policy be conducted by a nation's Congress or legislature comprised of elected representatives? Or should it be conducted by a politically appointed central bank that is more independent of voters? Here are some of the arguments made by each side.

The Case for Greater Democratic Control of Monetary Policy

Elected representatives conduct fiscal policy by passing tax and spending bills. They could handle monetary policy in the same way. Sure, they will sometimes make mistakes, but in a democracy, it is better to have mistakes made by elected officials accountable to voters than by political appointees. After all, the people appointed to the top governing positions at the Federal Reserve—and to most central banks around the world—are typically bankers

and economists. They are not representatives of borrowers like small businesses or farmers nor are they representatives of labor unions. Central banks might not be so quick to raise interest rates if they had to pay more attention to firms and people in the real economy.

The Case for an Independent Central Bank

Because the central bank has some insulation from day-to-day politics, its members can take a nonpartisan look at specific economic situations and make tough, immediate decisions when necessary. The idea of giving a legislature the ability to create money and hand out loans is likely to end up badly, sooner or later. It is simply too tempting for lawmakers to expand the money supply to fund their projects. The long term result will be rampant inflation. Also, a central bank, acting according to the laws passed by elected officials, can respond far more quickly than a legislature. For example, the U.S. budget takes months to debate, pass, and be signed into law, but monetary policy decisions can be made much more rapidly. Day-to-day democratic control of monetary policy is impractical and seems likely to lead to an overly expansionary monetary policy and higher inflation.

The problem of excess reserves does not affect contractionary policy. Central bankers have an old saying that monetary policy can be like pulling and pushing on a string: when the central bank pulls on the string and uses contractionary monetary policy, it can definitely raise interest rates and reduce aggregate demand. However, when the central bank tries to push on the string of expansionary monetary policy, the string may sometimes just fold up limp and have little effect, because banks decide not to loan out their excess reserves. This analogy should not be taken too literally—expansionary monetary policy usually does have real effects, after that inconveniently long and variable lag. There are also times, like Japan's economy in the late 1990s and early 2000s, when expansionary monetary policy has been insufficient to lift a recession-prone economy.

Unpredictable Movements of Velocity

Velocity is a term that economists use to describe how quickly money circulates through the economy. The **velocity** of money in a year is defined as:

Equation:

$$\text{Velocity} = \frac{\text{nominal GDP}}{\text{money supply}}$$

Specific measurements of velocity depend on the definition of the money supply being used. Consider the velocity of M1, the total amount of currency in circulation and checking account balances. In 2009, for example, M1 was \$1.7 trillion and nominal GDP was \$14.3 trillion, so the velocity of M1 was 8.4 (\$14.3 trillion/\$1.7 trillion). A higher velocity of money means that the average dollar circulates more times in a year; a lower velocity means that the average dollar circulates fewer times in a year.

Perhaps you heard the “d” word mentioned during our recent economic downturn. See the following Clear It Up feature for a discussion of how deflation could affect monetary policy.

Note:

What happens during episodes of deflation?

Deflation occurs when the rate of inflation is negative; that is, instead of money having less purchasing power over time, as occurs with inflation, money is worth more. Deflation can make it very difficult for monetary policy to address a recession.

Remember that the real interest rate is the nominal interest rate minus the rate of inflation. If the nominal interest rate is 7% and the rate of inflation is 3%, then the borrower is effectively paying a 4% real interest rate. If the nominal interest rate is 7% and there is *deflation* of 2%, then the real interest rate is actually 9%. In this way, an unexpected deflation raises the real interest payments for borrowers. It can lead to a situation where an unexpectedly high number of loans are not repaid, and banks find that their net worth is decreasing or negative. When banks are suffering losses, they become less able and eager to make new loans. Aggregate demand declines, which can lead to recession.

Then the double-whammy: After causing a recession, deflation can make it difficult for monetary policy to work. Say that the central bank uses expansionary monetary policy to reduce the nominal interest rate all the way

to zero—but the economy has 5% deflation. As a result, the real interest rate is 5%, and because a central bank cannot make the nominal interest rate negative, expansionary policy cannot reduce the real interest rate further. In the U.S. economy during the early 1930s, deflation was 6.7% per year from 1930–1933, which caused many borrowers to default on their loans and many banks to end up bankrupt, which in turn contributed substantially to the Great Depression. Not all episodes of deflation, however, end in economic depression. Japan, for example, experienced deflation of slightly less than 1% per year from 1999–2002, which hurt the Japanese economy, but it still grew by about 0.9% per year over this period. Indeed, there is at least one historical example of deflation coexisting with rapid growth. The U.S. economy experienced deflation of about 1.1% per year over the quarter-century from 1876–1900, but real GDP also expanded at a rapid clip of 4% per year over this time, despite some occasional severe recessions. The central bank should be on guard against deflation and, if necessary, use expansionary monetary policy to prevent any long-lasting or extreme deflation from occurring. Except in severe cases like the Great Depression, deflation does not guarantee economic disaster.

Changes in velocity can cause problems for monetary policy. To understand why, rewrite the definition of velocity so that the money supply is on the left-hand side of the equation. That is:

Equation:

$$\text{Money supply} \times \text{velocity} = \text{Nominal GDP}$$

Recall from [The Macroeconomic Perspective](#) that

Equation:

$$\text{Nominal GDP} = \text{Price Level (or GDP Deflator)} \times \text{Real GDP}.$$

Therefore,

Equation:

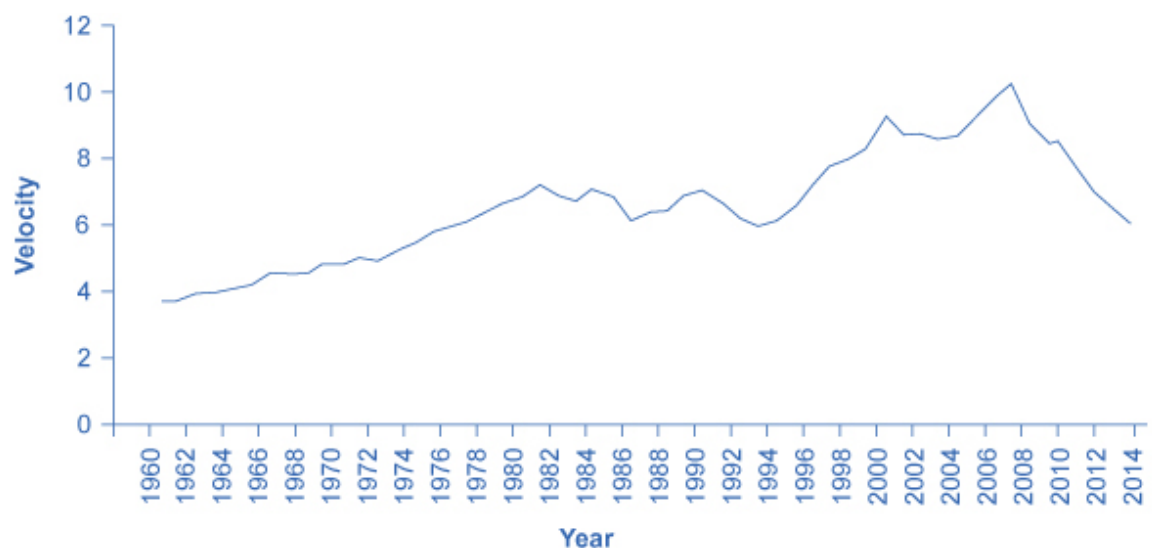
$$\text{Money Supply} \times \text{velocity} = \text{Nominal GDP} = \text{Price Level} \times \text{Real GDP}.$$

This equation is sometimes called the **basic quantity equation of money** but, as you can see, it is just the definition of velocity written in a different form. This equation must hold true, by definition.

If velocity is constant over time, then a certain percentage rise in the money supply on the left-hand side of the basic quantity equation of money will inevitably lead to the same percentage rise in nominal GDP—although this change could happen through an increase in inflation, or an increase in real GDP, or some combination of the two. If velocity is changing over time but in a constant and predictable way, then changes in the money supply will continue to have a predictable effect on nominal GDP. If velocity changes unpredictably over time, however, then the effect of changes in the money supply on nominal GDP becomes unpredictable.

The actual velocity of money in the U.S. economy as measured by using M1, the most common definition of the money supply, is illustrated in [\[link\]](#). From 1960 up to about 1980, velocity appears fairly predictable; that is, it is increasing at a fairly constant rate. In the early 1980s, however, velocity as calculated with M1 becomes more variable. The reasons for these sharp changes in velocity remain a puzzle. Economists suspect that the changes in velocity are related to innovations in banking and finance which have changed how money is used in making economic transactions: for example, the growth of electronic payments; a rise in personal borrowing and credit card usage; and accounts that make it easier for people to hold money in savings accounts, where it is counted as M2, right up to the moment that they want to write a check on the money and transfer it to M1. So far at least, it has proven difficult to draw clear links between these kinds of factors and the specific up-and-down fluctuations in M1. Given many changes in banking and the prevalence of electronic banking, M2 is now favored as a measure of money rather than the narrower M1.

Velocity Calculated Using M1



Velocity is the nominal GDP divided by the money supply for a given year. Different measures of velocity can be calculated by using different measures of the money supply. Velocity, as calculated by using M1, has lacked a steady trend since the 1980s, instead bouncing up and down. (credit: Federal Reserve Bank of St. Louis)

In the 1970s, when velocity as measured by M1 seemed predictable, a number of economists, led by Nobel laureate Milton Friedman (1912–2006), argued that the best monetary policy was for the central bank to increase the money supply at a constant growth rate. These economists argued that with the long and variable lags of monetary policy, and the political pressures on central bankers, central bank monetary policies were as likely to have undesirable as to have desirable effects. Thus, these economists believed that the monetary policy should seek steady growth in the money supply of 3% per year. They argued that a steady rate of monetary growth would be correct over longer time periods, since it would roughly match the growth of the real economy. In addition, they argued that giving the central bank less discretion to conduct monetary policy would prevent an overly activist central bank from becoming a source of economic instability and uncertainty. In this spirit, Friedman wrote in 1967: “The first and most important lesson that history teaches about what monetary policy can do—and it is a lesson of the most

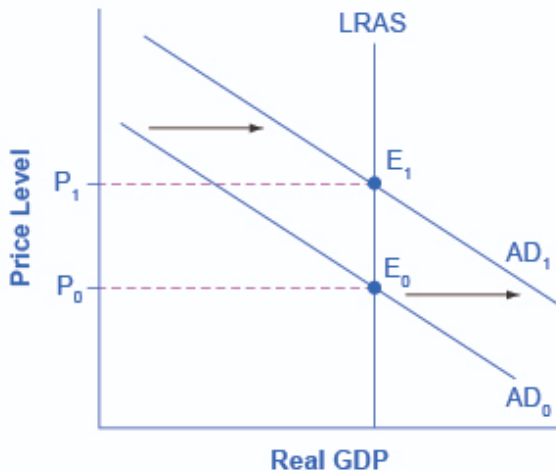
profound importance—is that monetary policy can prevent money itself from being a major source of economic disturbance.”

As the velocity of M1 began to fluctuate in the 1980s, having the money supply grow at a predetermined and unchanging rate seemed less desirable, because as the quantity theory of money shows, the combination of constant growth in the money supply and fluctuating velocity would cause nominal GDP to rise and fall in unpredictable ways. The jumpiness of velocity in the 1980s caused many central banks to focus less on the rate at which the quantity of money in the economy was increasing, and instead to set monetary policy by reacting to whether the economy was experiencing or in danger of higher inflation or unemployment.

Unemployment and Inflation

If you were to survey central bankers around the world and ask them what they believe should be the primary task of monetary policy, the most popular answer by far would be fighting inflation. Most central bankers believe that the neoclassical model of economics accurately represents the economy over the medium to long term. Remember that in the neoclassical model of the economy, the aggregate supply curve is drawn as a vertical line at the level of potential GDP, as shown in [\[link\]](#). In the neoclassical model, the level of potential GDP (and the natural rate of unemployment that exists when the economy is producing at potential GDP) is determined by real economic factors. If the original level of aggregate demand is AD_0 , then an expansionary monetary policy that shifts aggregate demand to AD_1 only creates an inflationary increase in the price level, but it does not alter GDP or unemployment. From this perspective, all that monetary policy can do is to lead to low inflation or high inflation—and low inflation provides a better climate for a healthy and growing economy. After all, low inflation means that businesses making investments can focus on real economic issues, not on figuring out ways to protect themselves from the costs and risks of inflation. In this way, a consistent pattern of low inflation can contribute to long-term growth.

Monetary Policy in a Neoclassical Model



In a neoclassical view, monetary policy affects only the price level, not the level of output in the economy. For example, an expansionary monetary policy causes aggregate demand to shift from the original AD_0 to AD_1 . However, the adjustment of the economy from the original equilibrium (E_0) to the new equilibrium (E_1) represents an inflationary increase in the price level from P_0 to P_1 , but has no effect in the long run on output or the unemployment rate. In fact, no shift in AD will affect the equilibrium quantity of output in this model.

This vision of focusing monetary policy on a low rate of inflation is so attractive that many countries have rewritten their central banking laws since in the 1990s to have their bank practice **inflation targeting**, which means that the central bank is legally required to focus primarily on keeping inflation low. By 2014, central banks in 28 countries, including Austria, Brazil, Canada, Israel, Korea, Mexico, New Zealand, Spain, Sweden, Thailand, and the United Kingdom faced a legal requirement to target the inflation rate. A notable exception is the Federal Reserve in the United States, which does not

practice inflation-targeting. Instead, the law governing the Federal Reserve requires it to take both unemployment and inflation into account.

Economists have no final consensus on whether a central bank should be required to focus only on inflation or should have greater discretion. For those who subscribe to the inflation targeting philosophy, the fear is that politicians who are worried about slow economic growth and unemployment will constantly pressure the central bank to conduct a loose monetary policy—even if the economy is already producing at potential GDP. In some countries, the central bank may lack the political power to resist such pressures, with the result of higher inflation, but no long-term reduction in unemployment. The U.S. Federal Reserve has a tradition of independence, but central banks in other countries may be under greater political pressure. For all of these reasons—long and variable lags, excess reserves, unstable velocity, and controversy over economic goals—monetary policy in the real world is often difficult. The basic message remains, however, that central banks can affect aggregate demand through the conduct of monetary policy and in that way influence macroeconomic outcomes.

Asset Bubbles and Leverage Cycles

One long-standing concern about having the central bank focus on inflation and unemployment is that it may be overlooking certain other economic problems that are coming in the future. For example, from 1994 to 2000 during what was known as the “dot-com” boom, the U.S. stock market, which is measured by the Dow Jones Industrial Index (which includes 30 very large companies from across the U.S. economy), nearly tripled in value. The Nasdaq index, which includes many smaller technology companies, increased in value by a multiple of five from 1994 to 2000. These rates of increase were clearly not sustainable. Indeed, stock values as measured by the Dow Jones were almost 20% lower in 2009 than they had been in 2000. Stock values in the Nasdaq index were 50% lower in 2009 than they had been in 2000. The drop-off in stock market values contributed to the recession of 2001 and the higher unemployment that followed.

A similar story can be told about housing prices in the mid-2000s. During the 1970s, 1980s, and 1990s, housing prices increased at about 6% per year on average. During what came to be known as the “housing bubble” from 2003

to 2005, housing prices increased at almost double this annual rate. These rates of increase were clearly not sustainable. When the price of housing fell in 2007 and 2008, many banks and households found that their assets were worth less than they expected, which contributed to the recession that started in 2007.

At a broader level, some economists worry about a leverage cycle, where “leverage” is a term used by financial economists to mean “borrowing.” When economic times are good, banks and the financial sector are eager to lend, and people and firms are eager to borrow. Remember that the amount of money and credit in an economy is determined by a money multiplier—a process of loans being made, money being deposited, and more loans being made. In good economic times, this surge of lending exaggerates the episode of economic growth. It can even be part of what lead prices of certain assets—like stock prices or housing prices—to rise at unsustainably high annual rates. At some point, when economic times turn bad, banks and the financial sector become much less willing to lend, and credit becomes expensive or unavailable to many potential borrowers. The sharp reduction in credit, perhaps combined with the deflating prices of a dot-com stock price bubble or a housing bubble, makes the economic downturn worse than it would otherwise be.

Thus, some economists have suggested that the central bank should not just look at economic growth, inflation, and unemployment rates, but should also keep an eye on asset prices and leverage cycles. Such proposals are quite controversial. If a central bank had announced in 1997 that stock prices were rising “too fast” or in 2004 that housing prices were rising “too fast,” and then taken action to hold down price increases, many people and their elected political representatives would have been outraged. Neither the Federal Reserve nor any other central banks want to take the responsibility of deciding when stock prices and housing prices are too high, too low, or just right. As further research explores how asset price bubbles and leverage cycles can affect an economy, central banks may need to think about whether they should conduct monetary policy in a way that would seek to moderate these effects.

Let’s end this chapter with a Work it Out exercise in how the Fed—or any central bank—would stir up the economy by increasing the money supply.

Note:**Calculating the Effects of Monetary Stimulus**

Suppose that the central bank wants to stimulate the economy by increasing the money supply. The bankers estimate that the velocity of money is 3, and that the price level will increase from 100 to 110 due to the stimulus. Using the quantity equation of money, what will be the impact of an \$800 billion dollar increase in the money supply on the quantity of goods and services in the economy given an initial money supply of \$4 trillion?

Step 1. We begin by writing the quantity equation of money: $MV = PQ$. We know that initially $V = 3$, $M = 4,000$ (billion) and $P = 100$. Substituting these numbers in, we can solve for Q :

Equation:

$$\begin{aligned}MV &= PQ \\4,000 \times 3 &= 100 \times Q \\Q &= 120\end{aligned}$$

Step 2. Now we want to find the effect of the addition \$800 billion in the money supply, together with the increase in the price level. The new equation is:

Equation:

$$\begin{aligned}MV &= PQ \\4,800 \times 3 &= 110 \times Q \\Q &= 130.9\end{aligned}$$

Step 3. If we take the difference between the two quantities, we find that the monetary stimulus increased the quantity of goods and services in the economy by 10.9 billion.

The discussion in this chapter has focused on domestic monetary policy; that is, the view of monetary policy within an economy. [Exchange Rates and International Capital Flows](#) explores the international dimension of monetary policy, and how monetary policy becomes involved with exchange rates and international flows of financial capital.

Note:**The Problem of the Zero Percent Interest Rate Lower Bound**

In 2008, the U.S. Federal Reserve found itself in a difficult position. The federal funds rate was on its way to near zero, which meant that traditional open market operations, by which the Fed purchases U.S. Treasury Bills to lower short term interest rates, was no longer viable. This so called “zero bound problem,” prompted the Fed, under then Chair Ben Bernanke, to attempt some unconventional policies, collectively called quantitative easing. By early 2014, quantitative easing nearly quintupled the amount of bank reserves. This likely contributed to the U.S. economy’s recovery, but the impact was muted, probably due to some of the hurdles mentioned in the last section of this module. The unprecedented increase in bank reserves also led to fears of inflation. As of early 2015, however, there have been no serious signs of a boom, with core inflation around a stable 1.7%.

Key Concepts and Summary

Monetary policy is inevitably imprecise, for a number of reasons: (a) the effects occur only after long and variable lags; (b) if banks decide to hold excess reserves, monetary policy cannot force them to lend; and (c) velocity may shift in unpredictable ways. The basic quantity equation of money is $MV = PQ$, where M is the money supply, V is the velocity of money, P is the price level, and Q is the real output of the economy. Some central banks, like the European Central Bank, practice inflation targeting, which means that the only goal of the central bank is to keep inflation within a low target range. Other central banks, such as the U.S. Federal Reserve, are free to focus on either reducing inflation or stimulating an economy that is in recession, whichever goal seems most important at the time.

Self-Check Questions**Exercise:****Problem:**

Why might banks want to hold excess reserves in time of recession?

Solution:

In times of economic uncertainty, banks may worry that borrowers will lose the ability to repay their loans. They may also fear that a panic is more likely and they will need the excess reserves to meet their obligations.

Exercise:

Problem: Why might the velocity of money change unexpectedly?

Solution:

If consumer optimism changes, spending can speed up or slow down. This could also happen in a case where consumers need to buy a large number of items quickly, such as in a situation of national emergency.

Review Questions**Exercise:****Problem:**

Which kind of monetary policy would you expect in response to recession: expansionary or contractionary? Why?

Exercise:**Problem:**

How might each of the following factors complicate the implementation of monetary policy: long and variable lags, excess reserves, and movements in velocity?

Exercise:

Problem: Define the velocity of the money supply.

Exercise:

Problem: What is the basic quantity equation of money?

Exercise:

Problem: How does a monetary policy of inflation targeting work?

Critical Thinking Questions

Exercise:

Problem:

How does rule-based monetary policy differ from discretionary monetary policy (that is, monetary policy not based on a rule)? What are some of the arguments for each?

Exercise:

Problem:

Is it preferable for central banks to primarily target inflation or unemployment? Why?

Problems

Exercise:

Problem:

All other things being equal, by how much will nominal GDP expand if the central bank increases the money supply by \$100 billion, and the velocity of money is 3? (Use this information as necessary to answer the following 4 questions.)

Exercise:

Problem:

Suppose now that economists expect the velocity of money to increase by 50% as a result of the monetary stimulus. What will be the total increase in nominal GDP?

Exercise:**Problem:**

If GDP is 1,500 and the money supply is 400, what is velocity?

Exercise:**Problem:**

If GDP now rises to 1,600, but the money supply does not change, how has velocity changed?

Exercise:**Problem:**

If GDP now falls back to 1,500 and the money supply falls to 350, what is velocity?

References

Tobin, James. "The Concise Encyclopedia of Economics: Monetary Policy." *Library of Economics and Liberty*. Accessed November 2013.
<http://www.econlib.org/library/Enc/MonetaryPolicy.html>.

Federal Reserve Bank of New York. "The Founding of the Fed." Accessed November 2013.
http://www.newyorkfed.org/aboutthefed/history_article.html.

Glossary

basic quantity equation of money

$$\text{money supply} \times \text{velocity} = \text{nominal GDP}$$

excess reserves

reserves banks hold that exceed the legally mandated limit

inflation targeting

a rule that the central bank is required to focus only on keeping inflation low

velocity

the speed with which money circulates through the economy; calculated as the nominal GDP divided by the money supply

Introduction to Poverty and Economic Inequality

class="introduction"

Occupying Wall Street

On September
17, 2011, Occupy
Wall Street began
in New York
City's Wall Street
financial district.

(Credit:
modification of
work by David
Shankbone/Flick
r Creative
Commons)



Note:
Occupy Wall Street

In September 2011, a group of protesters gathered in Zuccotti Park in New York City to decry what they perceived as increasing social and economic inequality in the United States. Calling their protest “Occupy Wall Street,” they argued that the concentration of wealth among the richest 1% in the United States was both economically unsustainable and inequitable, and needed to be changed. The protest then spread to other major cities, and the Occupy movement was born.

Why were people so upset? How much wealth is concentrated among the top 1% in our society? How did they acquire so much wealth? These are very real, very important questions in the United States now, and this chapter on poverty and economic inequality will help us address the causes behind this sentiment.

Note:

Introduction to Poverty and Economic Inequality

In this chapter, you will learn about:

- Drawing the Poverty Line
- The Poverty Trap
- The Safety Net
- Income Inequality: Measurement and Causes
- Government Policies to Reduce Income Inequality

The labor markets that determine what workers are paid do not take into account how much income a family needs for food, shelter, clothing, and health care. Market forces do not worry about what happens to families when a major local employer goes out of business. Market forces do not take time to contemplate whether those who are earning higher incomes should pay an even higher share of taxes.

However, labor markets do create considerable inequalities of income. In 2014, the median American family income was \$57,939 (the median is the level where half of all families had more than that level and half had less).

According to the U.S. Census Bureau, almost nine million U.S. families were classified by the federal government as being below the poverty line in that year. Think about a family of three—perhaps a single mother with two children—attempting to pay for the basics of life on perhaps \$17,916 per year. After paying for rent, healthcare, clothing, and transportation, such a family might have \$6,000 to spend on food. Spread over 365 days, the food budget for the entire family would be about \$17 per day. To put this in perspective, most cities have restaurants where \$17 will buy you an appetizer for one.

This chapter explores how the U.S. government defines poverty, the balance between assisting the poor without discouraging work, and how federal antipoverty programs work. It also discusses income inequality—how economists measure inequality, why inequality has changed in recent decades, the range of possible government policies to reduce inequality, and the danger of a tradeoff that too great a reduction in inequality may reduce incentives for producing output.

Drawing the Poverty Line

By the end of this section, you will be able to:

- Explain economic inequality and how the poverty line is determined
- Analyze the U.S. poverty rate over time, noting its prevalence among different groups of citizens

Comparisons of high and low incomes raise two different issues: economic inequality and **poverty**. Poverty is measured by the number of people who fall below a certain level of income—called the **poverty line**—that defines the income needed for a basic standard of living. **Income inequality** compares the share of the total income (or wealth) in society that is received by different groups; for example, comparing the share of income received by the top 10% to the share of income received by the bottom 10%.

In the United States, the official definition of the poverty line traces back to a single person: Mollie Orshansky. In 1963, Orshansky, who was working for the Social Security Administration, published an article called “Children of the Poor” in a highly useful and dry-as-dust publication called the *Social Security Bulletin*. Orshansky’s idea was to define a poverty line based on the cost of a healthy diet.

Her previous job had been at the U.S. Department of Agriculture, where she had worked in an agency called the Bureau of Home Economics and Human Nutrition. One task of this bureau had been to calculate how much it would cost to feed a nutritionally adequate diet to a family. Orshansky found that the average family spent one-third of its income on food. She then proposed that the poverty line be the amount needed to buy a nutritionally adequate diet, given the size of the family, multiplied by three.

The current U.S. poverty line is essentially the same as the Orshansky poverty line, although the dollar amounts are adjusted each year to represent the same buying power over time. The U.S. poverty line in 2015 ranged from \$11,790 for a single individual to \$25,240 for a household of four people.

[\[link\]](#) shows the U.S. **poverty rate** over time; that is, the percentage of the population below the poverty line in any given year. The poverty rate

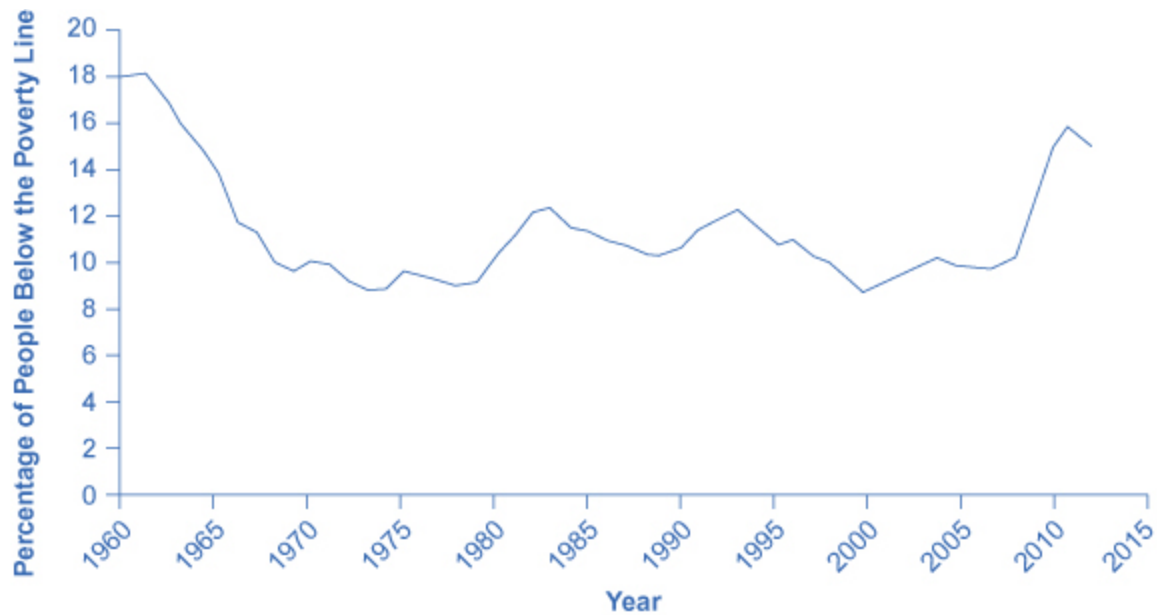
declined through the 1960s, rose in the early 1980s and early 1990s, but seems to have been slightly lower since the mid-1990s. However, in no year in the last four decades has the poverty rate been less than 11% of the U.S. population—that is, at best about one American in nine is below the poverty line. In recent years, the poverty rate appears to have peaked at 15.9% in 2011 before dropping to 14.5% in 2013. [\[link\]](#) compares poverty rates for different groups in 2011. As you will see when we delve further into these numbers, poverty rates are relatively low for whites, for the elderly, for the well-educated, and for male-headed households. Poverty rates for females, Hispanics, and African Americans are much higher than for whites. While Hispanics and African Americans have a higher percentage of individuals living in poverty than others, most people in the United States living below the poverty line are white.

Note:

Visit this [website](#) for more information on U.S. poverty.



The U.S. Poverty Rate since 1960



The poverty rate fell dramatically during the 1960s, rose in the early 1980s and early 1990s, and, after declining in the 1990s through mid-2000s, rose to 15.9% in 2011, which is close to the 1960 levels. In 2013, the poverty dropped slightly to 14.5%.
(Source: U.S. Census Bureau)

Group	Poverty Rate
Females	15.8%
Males	13.1%
White	9.6%

Group	Poverty Rate
Black	27.1%
Hispanic	23.5%
Under age 18	19.9%
Ages 18–24	20.6%
Ages 25–34	15.9%
Ages 35–44	12.2%
Ages 45–54	10.9%
Ages 55–59	10.7%
Ages 60–64	10.8%
Ages 65 and older	9.5%

Poverty Rates by Group, 2013

The concept of a poverty line raises many tricky questions. In a vast country like the United States, should there be a national poverty line? After all, according to the Federal Register, the median household income for a family of four was \$102,552 in New Jersey and \$57,132 in Mississippi in 2013, and prices of some basic goods like housing are quite different between states. The poverty line is based on cash income, which means it does not take into account government programs that provide assistance to the poor in a non-cash form, like Medicaid (health care for low-income individuals and families) and food aid. Also, low-income families can qualify for federal housing assistance. (These and other government aid programs will be discussed in detail later in this chapter.)

Should the poverty line be adjusted to take the value of such programs into account? Many economists and policymakers wonder whether the concept of what poverty means in the twenty-first century should be rethought. The following Clear It Up feature explains the poverty lines set by the World Bank for low-income countries around the world.

Note:

How is poverty measured in low-income countries?

The World Bank sets two poverty lines for low-income countries around the world. One poverty line is set at an income of \$1.25/day per person; the other is at \$2/day. By comparison, the U.S. 2015 poverty line of \$20,090 annually for a family of three works out to \$18.35 per person per day. Clearly, many people around the world are far poorer than Americans, as [\[link\]](#) shows. China and India both have more than a billion people; Nigeria is the most populous country in Africa; and Egypt is the most populous country in the Middle East. In all four of those countries, in the mid-2000s, a substantial share of the population subsisted on less than \$2/day. Indeed, about half the world lives on less than \$2.50 a day, and 80 percent of the world lives on less than \$10 per day. (Of course, the cost of food, clothing, and shelter in those countries can be very different from those costs in the United States, so the \$2 and \$2.50 figures may mean greater purchasing power than they would in the United States.)

Country	Share of Population below \$1.25/Day	Share of Population below \$2.00/Day
Brazil (in 2009)	6.1%	10.8%

Country	Share of Population below \$1.25/Day	Share of Population below \$2.00/Day
China (in 2009)	11.8%	27.2%
Egypt (in 2008)	1.7%	15.4%
India (in 2010)	32.7%	68.8%
Mexico (in 2010)	0.7%	4.5%
Nigeria (in 2010)	68.0%	84.5%
Poverty Lines for Low-Income Countries, mid-2000s(Source: http://data.worldbank.org/indicator/SI.POV.DDAY)		

Any poverty line will be somewhat arbitrary, and it is useful to have a poverty line whose basic definition does not change much over time. If Congress voted every few years to redefine what poverty means, then it would be difficult to compare rates over time. After all, would a lower poverty rate mean that the definition had been changed, or that people were actually better off? Government statisticians at the U.S. Census Bureau have ongoing research programs to address questions like these.

Key Concepts and Summary

Wages are influenced by supply and demand in labor markets, which can lead to very low incomes for some people and very high incomes for others. Poverty and income inequality are not the same thing. Poverty applies to the

condition of people who cannot afford the necessities of life. Income inequality refers to the disparity between those with higher and lower incomes. The poverty rate is what percentage of the population lives below the poverty line, which is determined by the amount of income that it takes to purchase the necessities of life. Choosing a poverty line will always be somewhat controversial.

Self-Check Questions

Exercise:

Problem:

Describe how each of these changes is likely to affect poverty and inequality:

- a. Incomes rise for low-income and high-income workers, but rise more for the high-income earners.
- b. Incomes fall for low-income and high-income workers, but fall more for high-income earners.

Solution:

- a. Poverty falls, inequality rises.
- b. Poverty rises, inequality falls.

Review Questions

Exercise:

Problem: How is the poverty rate calculated?

Exercise:

Problem: What is the poverty line?

Exercise:**Problem:**

What is the difference between poverty and income inequality?

Critical Thinking Questions**Exercise:****Problem:**

What goods and services would you include in an estimate of the basic necessities for a family of four?

Exercise:**Problem:**

If a family of three earned \$20,000, would they be able to make ends meet given the official poverty threshold?

Problems**Exercise:****Problem:**

In country A, the population is 300 million and 50 million people are living below the poverty line. What is the poverty rate?

Exercise:**Problem:**

In country B, the population is 900 million and 100 million people are living below the poverty line. What is the poverty rate?

References

Ebeling, Ashlea. 2014. "IRS Announces 2015 Estate And Gift Tax Limits." *Forbes*. Accessed March 16, 2015.
<http://www.forbes.com/sites/ashleaebeling/2014/10/30/irs-announces-2015-estate-and-gift-tax-limits/>.

Federal Register: The Daily Journal of the United States Government. "State Median Income Estimates for a Four-Person Household: Notice of the Federal Fiscal Year (FFY) 2013 State Median Income Estimates for Use Under the Low Income Home Energy Assistance Program (LIHEAP)." Last modified March 15, 2012.
<https://www.federalregister.gov/articles/2012/03/15/2012-6220/state-median-income-estimates-for-a-four-person-household-notice-of-the-federal-fiscal-year-ffy-2013#t-1>.

Luhby, Tami. 2014. "Income is on the rise . . . finally!" Accessed April 10, 2015. <http://money.cnn.com/2014/08/20/news/economy/median-income/>.

Meyer, Ali. 2015. "56,023,000: Record Number of Women Not in Labor Force." *CNSNews.com*. Accessed March 16, 2015.
<http://cnsnews.com/news/article/ali-meyer/56023000-record-number-women-not-labor-force>.

Orshansky, Mollie. "Children of the Poor." *Social Security Bulletin*. 26 no. 7 (1963): 3–13. <http://www.ssa.gov/policy/docs/ssb/v26n7/v26n7p3.pdf>.

The World Bank. "Data: Poverty Headcount Ratio at \$1.25 a Day (PPP) (% of Population)." <http://data.worldbank.org/indicator/SI.POV.DDAY>.

U.S. Department of Commerce: United States Census Bureau. "American FactFinder." <http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml>.

U.S. Department of Commerce: United States Census Bureau. "Current Population Survey (CPS): CPS Table Creator."
<http://www.census.gov/cps/data/cpstablecreator.html>.

U.S. Department of Commerce: United States Census Bureau. "Income: Table F-6. Regions—Families (All Races) by Median and Mean Income." <http://www.census.gov/hhes/www/income/data/historical/families/>.

U.S. Department of Commerce: United States Census Bureau. "Poverty: Poverty Thresholds." Last modified 2012.
<http://www.census.gov/hhes/www/poverty/data/threshld/>.

U.S. Department of Health & Human Services. 2015. "2015 Poverty Guidelines." Accessed April 10, 2015. <http://www.medicaid.gov/medicaid-chip-program-information/by-topics/eligibility/downloads/2015-federal-poverty-level-charts.pdf>.

U.S. Department of Health & Human Services. 2015. "Information on Poverty and Income Statistics: A Summary of 2014 Current Population Survey Data." Accessed April 13, 2015.
http://aspe.hhs.gov/hsp/14/povertyandincomeest/ib_poverty2014.pdf.

Glossary

income inequality

when one group receives a disproportionate share of total income or wealth than others

poverty

the situation of being below a certain level of income needed for a basic standard of living

poverty line

the specific amount of income needed for a basic standard of living

poverty rate

percentage of the population living below the poverty line

The Poverty Trap

By the end of this section, you will be able to:

- Explain the poverty trap, noting how it is impacted by government programs
- Identify potential issues in government programs that seek to reduce poverty
- Calculate a budget constraint line that represents the poverty trap

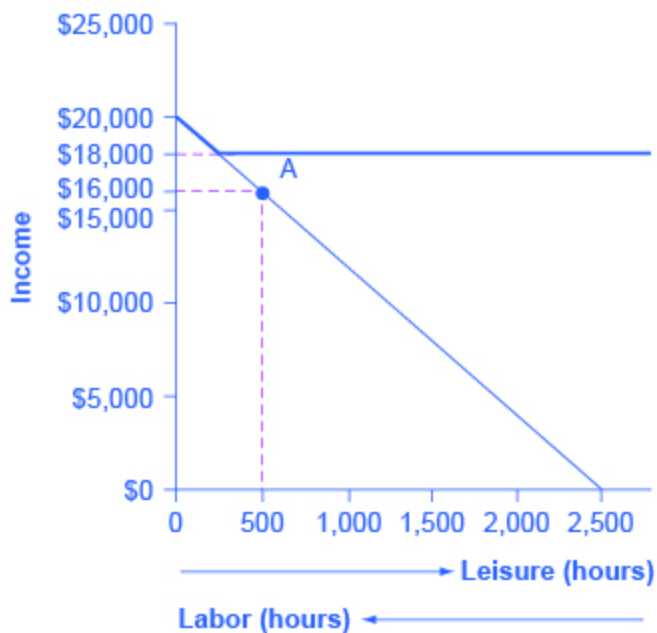
Can you give people too much help, or the wrong kind of help? When people are provided with food, shelter, healthcare, income, and other necessities, assistance may reduce their incentive to work. Consider a program to fight poverty that works in this reasonable-sounding manner: the government provides assistance to the poor, but as the poor earn income to support themselves, the government reduces the level of assistance it provides. With such a program, every time a poor person earns \$100, the person loses \$100 in government support. As a result, the person experiences no net gain for working. Economists call this problem the **poverty trap**.

Consider the situation faced by a single-parent family. A single mother (earning \$8 an hour) with two children, as illustrated in [\[link\]](#). First, consider the labor-leisure budget constraint faced by this family in a situation without government assistance. On the horizontal axis is hours of leisure (or time spent with family responsibilities) increasing in quantity from right to left. Also on the horizontal axis is the number of hours at paid work, going from zero hours on the right to the maximum of 2,500 hours on the left. On the vertical axis is the amount of income per year rising from low to higher amounts of income. The budget constraint line shows that at zero hours of leisure and 2,500 hours of work, the maximum amount of income is \$20,000 ($\$8 \times 2,500$ hours). At the other extreme of the budget constraint line, an individual would work zero hours, earn zero income, but enjoy 2,500 hours of leisure. At point A on the budget constraint line, by working 40 hours a week, 50 weeks a year, the utility-maximizing choice is to work a total of 2,000 hours per year and earn \$16,000.

Now suppose that a government antipoverty program guarantees every family with a single mother and two children \$18,000 in income. This is

represented on the graph by a horizontal line at \$18,000. With this program, each time the mother earns \$1,000, the government will deduct \$1,000 of its support. [\[link\]](#) shows what will happen at each combination of work and government support.

The Poverty Trap in Action



The original choice is 500 hours of leisure, 2,000 hours of work at point A, and income of \$16,000. With a guaranteed income of \$18,000, this family would receive \$18,000 whether it provides zero hours of work or 2,000 hours of work. Only if the family provides, say, 2,300 hours of work does its income rise above the guaranteed level of \$18,000—and even then, the marginal gain to income from working many hours is small.

Amount Worked (hours)	Total Earnings	Government Support	Total Income
0	0	\$18,000	\$18,000
500	\$4,000	\$14,000	\$18,000
1,000	\$8,000	\$10,000	\$18,000
1,500	\$12,000	\$6,000	\$18,000
2,000	\$16,000	\$2,000	\$18,000
2,500	\$20,000	0	\$20,000

Total Income at Various Combinations of Work and Support

The new budget line, with the antipoverty program in place, is the horizontal and heavy line that is flat at \$18,000. If the mother does not work at all, she receives \$18,000, all from the government. If she works full time, giving up 40 hours per week with her children, she still ends up with \$18,000 at the end of the year. Only if she works 2,300 hours in the year—which is an average of 44 hours per week for 50 weeks a year—does household income rise to \$18,400. Even in this case, all of her year's work means that household income rises by only \$400 over the income she would receive if she did not work at all. She would need to work 50 hours a week to reach \$20,000.

Indeed, the poverty trap is even stronger than this simplified example shows, because a working mother will have extra expenses like clothing, transportation, and child care that a nonworking mother will not face, making the economic gains from working even smaller. Moreover, those who do not work fail to build up job experience and contacts, which makes working in the future even less likely.

The bite of the poverty trap can be reduced by designing an antipoverty program so that, instead of reducing government payments by \$1 for every

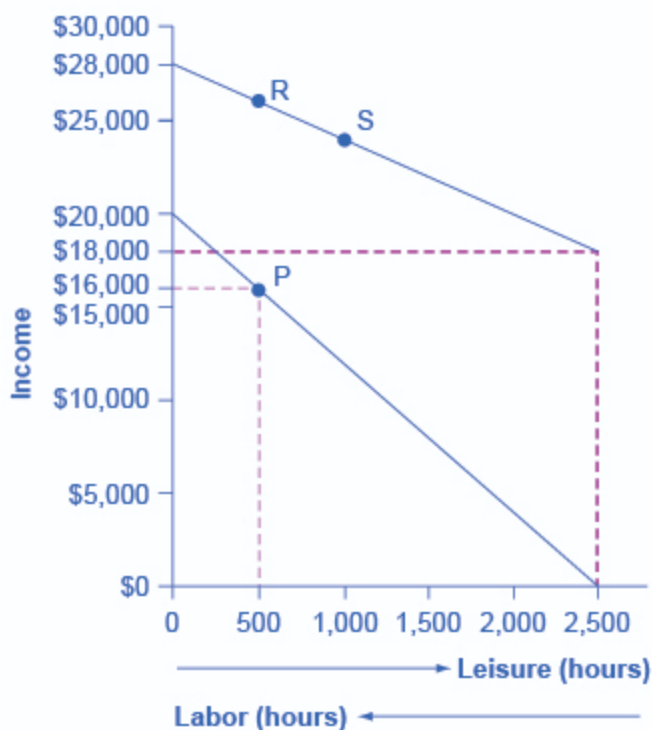
\$1 earned, payments are reduced by some smaller amount instead. The bite of the poverty trap can also be reduced by imposing requirements for work as a condition of receiving benefits and setting a time limit on benefits.

[\[link\]](#) illustrates a government program that guarantees \$18,000 in income, even for those who do not work at all, but then reduces this amount by 50 cents for each \$1 earned. The new, higher budget line in [\[link\]](#) shows that, with this program, additional hours of work will bring some economic gain. Because of the reduction in government income when an individual works, an individual earning \$8.00 will really net only \$4.00 per hour. The vertical intercept of this higher budget constraint line is at \$28,000 ($\$18,000 + 2,500 \text{ hours} \times \$4.00 = \$28,000$). The horizontal intercept is at the point on the graph where \$18,000 and 2500 hours of leisure is set. [\[link\]](#) shows the total income differences with various choices of labor and leisure.

However, this type of program raises other issues. First, even if it does not eliminate the incentive to work by reducing government payments by \$1 for every \$1 earned, enacting such a program may still reduce the incentive to work. At least some people who would be working 2,000 hours each year without this program might decide to work fewer hours but still end up with more income—that is, their choice on the new budget line would be like S, above and to the right of the original choice P. Of course, others may choose a point like R, which involves the same amount of work as P, or even a point to the left of R that involves more work.

The second major issue is that when the government phases out its support payments more slowly, the antipoverty program costs more money. Still, it may be preferable in the long run to spend more money on a program that retains a greater incentive to work, rather than spending less money on a program that nearly eliminates any gains from working.

Loosening the Poverty Trap: Reducing Government Assistance by 50 Cents for Every \$1 Earned



On the original labor-leisure opportunity set, the lower budget set shown by the smaller dashed line in the figure, the preferred choice P is 500 hours of leisure and \$16,000 of income. Then, the government created an antipoverty program that guarantees \$18,000 in income even to those who work zero hours, shown by the larger dashed line. In addition, every \$1 earned means phasing out 50 cents of benefits. This program leads to the higher budget set shown in the diagram. The hope is that this program will provide incentives to work the same or more hours, despite receiving income assistance. However, it is possible that the recipients will choose a point on the new budget set like S, with less work, more leisure, and greater

income, or a point like R, with the same work and greater income.

Amount Worked (hours)	Total Earnings	Government Support	Total Income
0	0	\$18,000	\$18,000
500	\$4,000	\$16,000	\$20,000
1,000	\$8,000	\$14,000	\$22,000
1,500	\$12,000	\$12,000	\$24,000
2,000	\$16,000	\$10,000	\$26,000
2,500	\$20,000	\$8,000	\$28,000

The Labor-Leisure Tradeoff with Assistance Reduced by 50 Cents for Every Dollar Earned

The next module will consider a variety of government support programs focused specifically on the poor, including welfare, SNAP (food supplement), Medicaid, and the earned income tax credit (EITC). Although these programs vary from state to state, it is generally a true statement that in many states from the 1960s into the 1980s, if poor people worked, their level of income barely rose—or did not rise at all—after the reduction in government support payments was factored in. The following Work It Out feature shows how this happens.

Note:**Calculating a Budget Constraint Line**

Jason earns \$9.00 an hour, and a government antipoverty program provides a floor of \$10,000 guaranteed income. The government reduces government support by \$0.50 for each \$1.00 earned. What are the horizontal and vertical intercepts of the budget constraint line? Assume the maximum hours for work or leisure is 2,500 hours.

Step 1. Determine the amount of the government guaranteed income. In this case, it is \$10,000.

Step 2. Plot that guaranteed income as a horizontal line on the budget constraint line.

Step 3. Determine what Jason earns if he has no income and enjoys 2,500 hours of leisure. In this case, he will receive the guaranteed \$10,000 (the horizontal intercept).

Step 4. Calculate how much Jason's salary will be reduced by due to the reduction in government income. In Jason's case, it will be reduced by one half. He will, in effect, net only \$4.50 an hour.

Step 5. If Jason works 1,000 hours, at a maximum what income will Jason receive? Jason will get the government assistance of \$10,000. He will net only \$4.50 for every hour he chooses to work. If he works 1,000 hours at \$4.50, his earned income is \$4,500 plus the government income of \$10,000. Thus the total maximum income (the vertical intercept) is $\$10,000 + \$4,500 = \$14,500$.

Key Concepts and Summary

A poverty trap occurs when government-support payments for the poor decline as the poor earn more income. As a result, the poor do not end up with much more income when they work, because the loss of government support largely or completely offsets any income that is earned by working. The bite of the poverty trap can be reduced by phasing out government benefits more slowly, as well as by imposing requirements for work as a condition of receiving benefits and a time limit on benefits.

Self-Check Questions

Exercise:**Problem:**

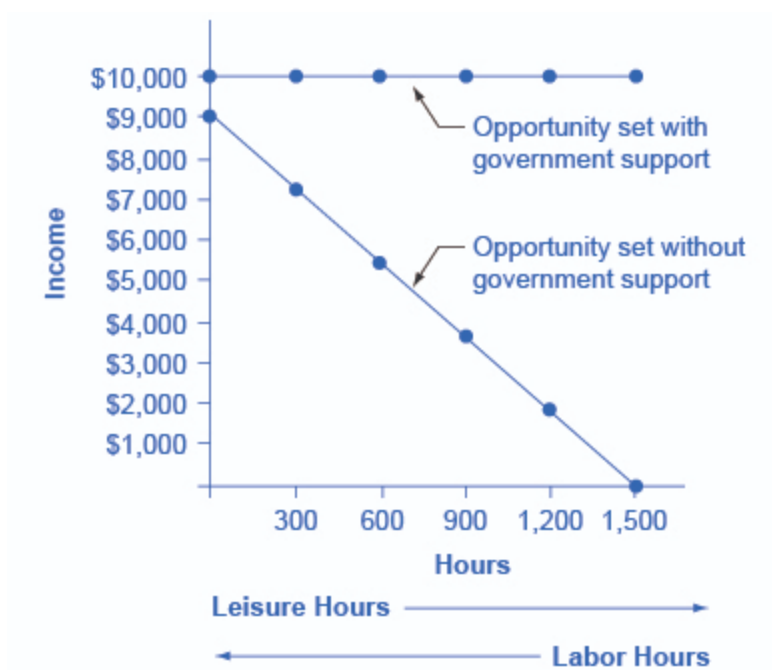
Jonathon is a single father with one child. He can work as a server for \$6 per hour for up to 1,500 hours per year. He is eligible for welfare, and so if he does not earn any income, he will receive a total of \$10,000 per year. He can work and still receive government benefits, but for every \$1 of income, his welfare stipend is \$1 less. Create a table similar to [\[link\]](#) that shows Jonathan's options. Use four columns, the first showing number of hours to work, the second showing his earnings from work, the third showing the government benefits he will receive, and the fourth column showing his total income (earnings + government support). Sketch a labor-leisure diagram of Jonathan's opportunity set with and without government support.

Solution:

Jonathon's options for working and total income are shown in the following table. His labor-leisure diagram is shown in the figure following the table.

Number of Work Hours	Earnings from Work	Government Benefits	Total Income
1,500	\$9,000	\$1,000	\$10,000
1,200	\$7,200	\$2,800	\$10,000
900	\$5,400	\$4,600	\$10,000
600	\$3,600	\$6,400	\$10,000

Number of Work Hours	Earnings from Work	Government Benefits	Total Income
300	\$1,800	\$8,200	\$10,000
0	\$0	\$10,000	\$10,000



Exercise:

Problem:

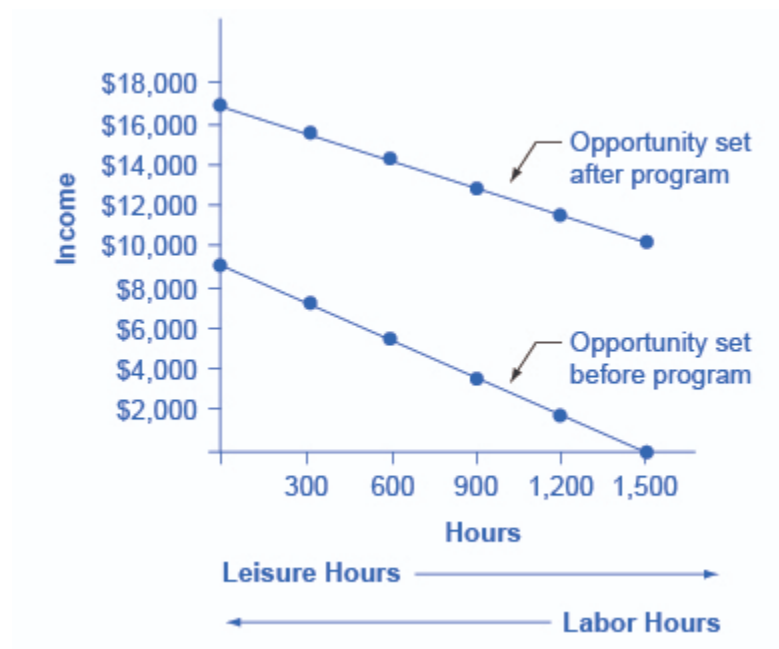
Imagine that the government reworks the welfare policy that was affecting Jonathan in question 1, so that for each dollar someone like Jonathan earns at work, his government benefits diminish by only 30 cents. Reconstruct the table from question 1 to account for this change in policy. Draw Jonathan's labor-leisure opportunity sets, both for before this welfare program is enacted and after it is enacted.

Solution:

The following table shows a policy where only 30 cents in government support is pulled right back for every \$1 of income earned. Jonathon's labor-leisure diagram is shown in the figure following the table.

“Opportunity set after program” extends from (0, \$16,300) to (1,500, \$10,000). “Opportunity set before program” slopes downward from (0, \$9,000) to (1,500, \$0).

Number of Work Hours	Earnings from Work	Government Benefits	Total Income
1,500	\$9,000	\$7,300	\$16,300
1,200	\$7,200	\$7,840	\$15,040
900	\$5,400	\$8,380	\$13,780
600	\$3,600	\$8,920	\$12,520
300	\$1,800	\$9,460	\$22,260
0	\$0	\$10,000	\$10,000



Review Questions

Exercise:

Problem:

How does the poverty trap discourage people from working?

Exercise:

Problem: How can the effect of the poverty trap be reduced?

Critical Thinking Questions

Exercise:

Problem:

[\[link\]](#) and [\[link\]](#) asked you to describe the labor-leisure tradeoff for Jonathon. Since, in the first example, there is no monetary incentive for Jonathon to work, explain why he may choose to work anyway. Explain what the opportunity costs of working and not working might be for Jonathon in each example. Using your tables and graphs from [\[link\]](#) and [\[link\]](#), analyze how the government welfare system affects Jonathan's incentive to work.

Exercise:

Problem:

Explain how you would create a government program that would give an incentive for labor to increase hours and keep labor from falling into the poverty trap.

Problems

Exercise:

Problem:

Susan is a single mother with three children. She can earn \$8 per hour and works up to 2,000 hours per year. However, if she does not earn any income at all, she will receive government benefits totaling \$16,000 per year. For every \$1 of income she earns, her level of government support will be reduced by \$1. Create a table, patterned after [\[link\]](#). The first column should show Susan's choices of how many hours to work per year, up to 2,000 hours. The second column should show her earnings from work. The third column should show her level of government support, given her earnings. The final column should show her total income, combining earnings and government support.

Glossary

poverty trap

antipoverty programs set up so that government benefits decline substantially as people earn more income—as a result, working provides little financial gain

The Safety Net

By the end of this section, you will be able to:

- Identify the antipoverty government programs that compose the safety net
- Explain the primary goals of the safety net programs and how these programs have changed over time
- Discuss the complexities of these safety net programs and why they can be controversial

The U.S. government has implemented a number of programs to assist those below the poverty line and those who have incomes just above the poverty line, who are referred to as the **near-poor**. Such programs are called the **safety net**, in recognition of the fact that they offer some protection for those who find themselves without jobs or income.

Temporary Assistance for Needy Families

From the Great Depression of the 1930s until 1996, the United States' most visible antipoverty program was Aid to Families with Dependent Children (AFDC), which provided cash payments to mothers with children who were below the poverty line. This program was often just called “welfare.” In 1996, Congress passed and President Bill Clinton signed into law the Personal Responsibility and Work Opportunity Reconciliation Act, more commonly called the “welfare reform act.” The new law replaced AFDC with Temporary Assistance for Needy Families (TANF).

Note:

Visit this [website](#) to watch a video of President Bill Clinton's Welfare Reform speech.



TANF brought several dramatic changes in how welfare operated. Under the old AFDC program, states set the level of welfare benefits that they would pay to the

poor, and the federal government guaranteed it would chip in some of the money as well. The federal government's welfare spending would rise or fall depending on the number of poor people, and on how each state set its own welfare contribution.

Under TANF, however, the federal government gives a fixed amount of money to each state. The state can then use the money for almost any program with an antipoverty component: for example, the state might use the money to give cash to poor families, or to reduce teenage pregnancy, or even to raise the high school graduation rate. However, the federal government imposed two key requirements. First, if states are to keep receiving the TANF grants, they must impose work requirements so that most of those receiving TANF benefits are working (or attending school). Second, no one can receive TANF benefits with federal money for more than a total of five years over his or her lifetime. The old AFDC program had no such work requirements or time limits.

TANF attempts to avoid the poverty trap by requiring that welfare recipients work and by limiting the length of time they can receive benefits. In its first few years, the program was quite successful. The number of families receiving payments in 1995, the last year of AFDC, was 4.8 million. By 2012, according to the Congressional Research Service, the average number of families receiving payments under TANF was 1.8 million—a decline of more than half.

TANF benefits to poor families vary considerably across states. For example, again according to the Congressional Research Service, in 2011 the highest monthly payment in Alaska to a single mother with two children was \$923, while in Mississippi the highest monthly payment to that family was \$170. These payments reflect differences in states' cost of living. Total spending on TANF was approximately \$16.6 billion in 1997. As of 2012, spending was at \$12 billion, an almost 28% decrease, split about evenly between the federal and state governments. When you take into account the effects of inflation, the decline is even greater. Moreover, there seemed little evidence that poor families were suffering a reduced standard of living as a result of TANF—although, on the other side, there was not much evidence that poor families had greatly improved their total levels of income, either.

The Earned Income Tax Credit (EITC)

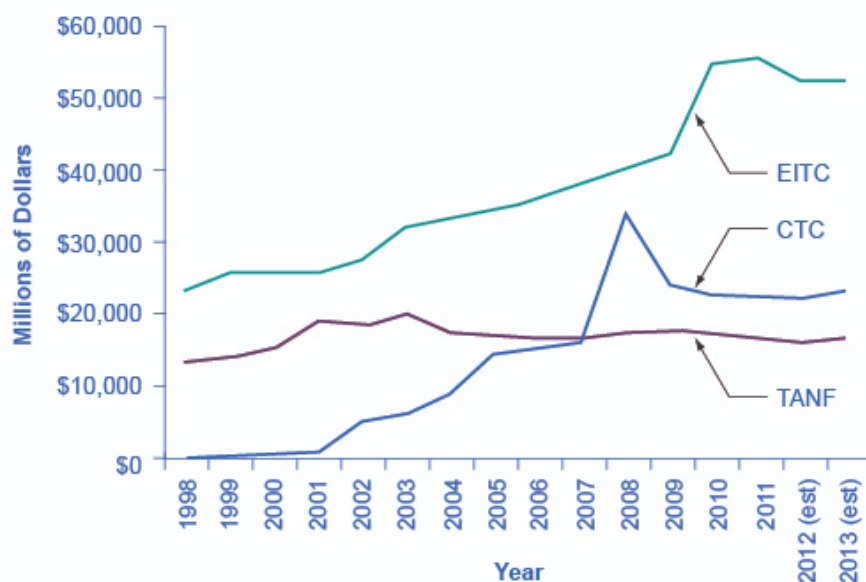
The **earned income tax credit (EITC)**, first passed in 1975, is a method of assisting the working poor through the tax system. The EITC is one of the largest assistance program for low-income groups, and projections for 2013 expected 26 million households to take advantage of it at an estimated cost of \$50 billion. In 2013, for example, a single parent with two children would have received a tax credit of \$5,372

up to an income level of \$17,530. The amount of the tax break increases with the amount of income earned, up to a point. The earned income tax credit has often been popular with both economists and the general public because of the way it effectively increases the payment received for work.

What about the danger of the poverty trap that every additional \$1 earned will reduce government support payments by close to \$1? To minimize this problem, the earned income tax credit is phased out slowly. According to the Tax Policy Center, for a single-parent family with two children in 2013, the credit is not reduced at all (but neither is it increased) as earnings rise from \$13,430 to \$17,530. Then, for every \$1 earned above \$17,530, the amount received from the credit is reduced by 21.06 cents, until the credit phases out completely at an income level of \$46,227.

[\[link\]](#) illustrates that the earned income tax credits, child tax credits, and the TANF program all cost the federal government money—either in direct outlays or in loss of tax revenues. CTC stands for the government tax cuts for the child tax credit.

Real Federal Spending on CTC, EITC, and TANF, 1975-2013



EITC increased from more than \$20 billion in 2000 to over an estimated \$50 billion by 2013, far exceeding estimated 2013 outlays in the CTC (Child Tax Credits) and TANF of over \$20 billion and \$10 billion, respectively. (Source: Office of Management and Budget)

In recent years, the EITC has become a hugely expensive government program for providing income assistance to the poor and near-poor, costing about \$60 billion in 2012. In that year, the EITC provided benefits to about 27 million families and individuals and, on average, is worth about \$2,296 per family (with children), according to the Tax Policy Center. One reason that the TANF law worked as well as it did is that the EITC was greatly expanded in the late 1980s and again in the early 1990s, which increased the returns to work for low-income Americans.

Supplemental Nutrition Assistance Program (SNAP)

Often called “food stamps,” **Supplemental Nutrition Assistance Program (SNAP)** is a federally funded program, started in 1964, in which each month poor people receive a card like a debit card that they can use to buy food. The amount of food aid for which a household is eligible varies by income, number of children, and other factors but, in general, households are expected to spend about 30% of their own net income on food, and if 30% of their net income is not enough to purchase a nutritionally adequate diet, then those households are eligible for SNAP.

SNAP can contribute to the poverty trap. For every \$100 earned, the government assumes that a family can spend \$30 more for food, and thus reduces its eligibility for food aid by \$30. This decreased benefit is not a complete disincentive to work—but combined with how other programs reduce benefits as income increases, it adds to the problem. SNAP, however, does try to address the poverty trap with its own set of work requirements and time limits.

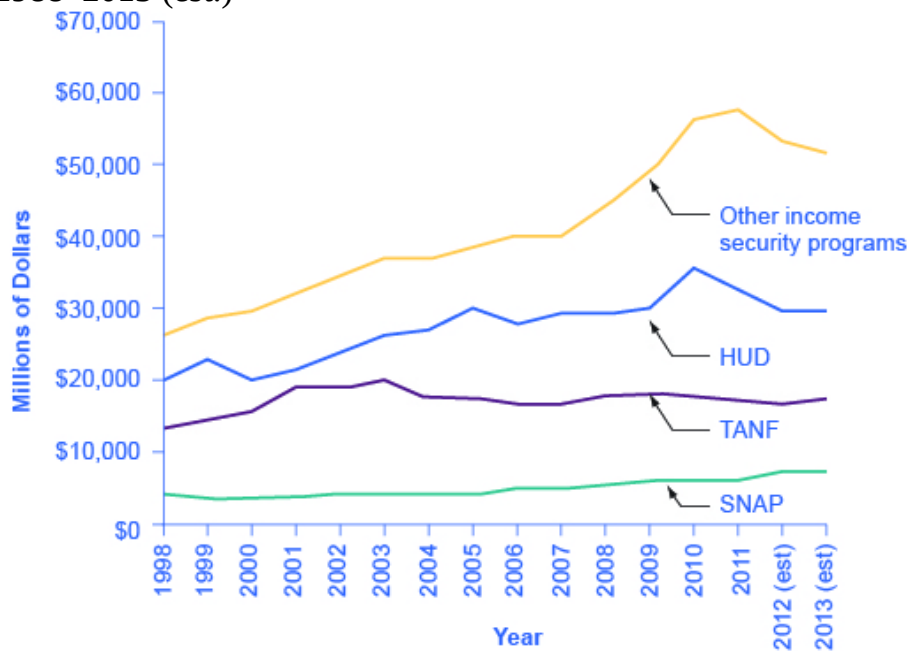
Why give debit cards and not just cash? Part of the political support for SNAP comes from a belief that since the cards must be spent on food, they cannot be “wasted” on other forms of consumption. From an economic point of view, however, the belief that cards must increase spending on food seems wrong-headed. After all, say that a poor family is spending \$2,500 per year on food, and then it starts receiving \$1,000 per year in SNAP aid. The family might react by spending \$3,500 per year on food (income plus aid), or it might react by continuing to spend \$2,500 per year on food, but use the \$1,000 in food aid to free up \$1,000 that can now be spent on other goods. So it is reasonable to think of SNAP cards as an alternative method, along with TANF and the earned income tax credit, of transferring income to the working poor.

Indeed, anyone eligible for TANF is also eligible for SNAP, although states can expand eligibility for food aid if they wish to do so. In some states, where TANF welfare spending is relatively low, a poor family may receive more in support from SNAP than from TANF. In 2014, about 40 million people received food aid at an annual cost of about \$76 billion, with an average monthly benefit of about \$287 per person per month. SNAP participation increased by 70% between 2007 and 2011,

from 26.6 million participants to 45 million. According to the Congressional Budget Office, this dramatic rise in participation was caused by the Great Recession of 2008–2009 and rising food prices.

The federal government deploys a range of income security programs that are funded through departments such as Health and Human Services, Agriculture, and Housing and Urban Development (HUD) (see [\[link\]](#)). According to the Office of Management and Budget, collectively, these three departments provided an estimated \$62 billion of aid through programs such as supplemental feeding programs for women and children, subsidized housing, and energy assistance. The federal government also transfers funds to individual states through special grant programs.

Expenditure Comparison of TANF, SNAP, HUD, and Other Income Security Programs, 1988–2013 (est.)



Total expenditures on income security continued to rise between 1988 and 2010, while payments for TANF have increased from \$13 billion in 1998 to an estimated \$17.3 billion in 2013. SNAP has seen relatively small increments. These two programs comprise a relatively small portion of the estimated \$106 billion dedicated to income security in 2013. Note that other programs and housing programs increased dramatically during the 2008 and 2010 time periods.

(Source: Table 12.3 Section 600 Income Security,
<https://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/hist.pdf>
)

The safety net includes a number of other programs: government-subsidized school lunches and breakfasts for children from low-income families; the Special Supplemental Food Program for Women, Infants and Children (WIC), which provides food assistance for pregnant women and newborns; the Low Income Home Energy Assistance Program, which provides help with home heating bills; housing assistance, which helps pay the rent; and Supplemental Security Income, which provides cash support for the disabled and the elderly poor.

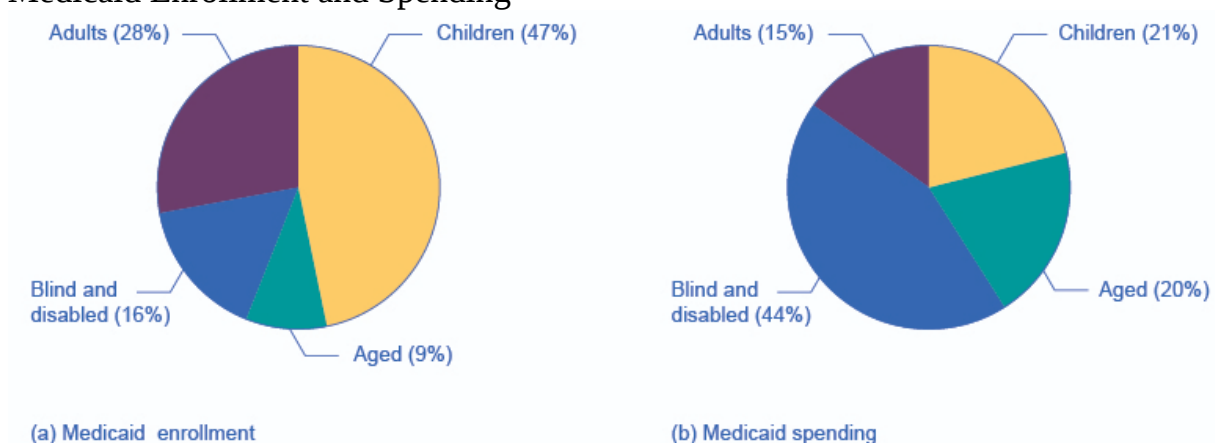
Medicaid

Medicaid was created by Congress in 1965 and is a joint health insurance program entered into by both the states and the federal government. The federal government helps fund Medicaid, but each state is responsible for administering the program, determining the level of benefits, and determining eligibility. It provides medical insurance for certain low-income people, including those below the poverty line, with a focus on families with children, the elderly, and the disabled. About one-third of Medicaid spending is for low-income mothers with children. While an increasing share of the program funding in recent years has gone to pay for nursing home costs for the elderly poor. The program ensures that a basic level of benefits is provided to Medicaid participants, but because each state sets eligibility requirements and provides varying levels of service, the program differs from state to state.

In the past, a common problem has been that many low-paying jobs pay enough to a breadwinner so that a family could lose its eligibility for Medicaid, yet the job does not offer health insurance benefits. A poor parent considering such a job might choose not to work rather than lose health insurance for his or her children. In this way, health insurance can become a part of the poverty trap. Many states recognized this problem in the 1980s and 1990s and expanded their Medicaid coverage to include not just the poor, but the near-poor earning up to 135% or even 185% of the poverty line. Some states also guaranteed that children would not lose coverage if their parents worked.

These expanded guarantees cost the government money, of course, but they also helped to encourage those on welfare to enter the labor force. As of 2014, approximately 69.7 million people participated in Medicaid. Of those enrolled, almost half are children. Healthcare expenditures, however, are highest for the elderly population, which comprises approximately 25% of participants. As [\[link\]](#) (a) indicates, the largest number of households that enroll in Medicaid are those with children. Lower-income adults are the next largest group enrolled in Medicaid at 28%. The blind and disabled are 16% of those enrolled, and seniors are 9% of those enrolled. [\[link\]](#) (b) shows how much actual Medicaid dollars are spent for each group. Out of total Medicaid spending, more is spent on seniors (20%) and the blind and

disabled (44%). So, 64% of all Medicaid spending goes to seniors, the blind, and disabled. Children receive 21% of all Medicaid spending, followed by adults at 15%. **Medicaid Enrollment and Spending**



Part (a) shows the Medicaid enrollment by different populations, with children comprising the largest percentage at 47%, followed by adults at 28%, and the blind and disabled at 16%. Part (b) shows that Medicaid spending is principally for the blind and disabled, followed by the elderly. Although children are the largest population covered by Medicaid, expenditures on children are only at 21%.

Key Concepts and Summary

The group of government programs that assist the poor are called the safety net. In the United States, prominent safety net programs include Temporary Assistance to Needy Families (TANF), the Supplemental Nutrition Assistance Program (SNAP), the earned income tax credit (EITC), Medicaid, and the Special Supplemental Food Program for Women, Infants, and Children (WIC).

Self-Check Questions

Exercise:

Problem:

We have discovered that the welfare system discourages recipients from working because the more income they earn, the less welfare benefits they receive. How does the earned income tax credit attempt to loosen the poverty trap?

Solution:

The earned income tax credit works like this: a poor family receives a tax break that increases according to how much they work. Families that work more get more. In that sense it loosens the poverty trap by encouraging work. As families earn above the poverty level, the earned income tax credit is gradually reduced. For those near-poor families, the earned income tax credit is a partial disincentive to work.

Exercise:

Problem: How does the TANF attempt to loosen the poverty trap?

Solution:

TANF attempts to loosen the poverty trap by providing incentives to work in other ways. Specifically, it requires that people work (or complete their education) as a condition of receiving TANF benefits, and it places a time limit on benefits.

Review Questions**Exercise:**

Problem: Who are the near-poor?

Exercise:

Problem: What is the safety net?

Exercise:**Problem:**

Briefly explain the differences between TANF, the earned income tax credit, SNAP, and Medicaid.

Critical Thinking Questions**Exercise:**

Problem:

Many critics of government programs to help low-income individuals argue that these programs create a poverty trap. Explain how programs such as TANF, EITC, SNAP, and Medicaid will affect low-income individuals and whether or not you think these programs will benefit families and children.

Exercise:**Problem:**

Think about the business cycle: during a recession, unemployment increases; it decreases in an expansionary phase. Explain what happens to TANF, SNAP, and Medicaid programs at each phase of the business cycle (recession, trough, expansion, and peak).

References

Congressional Budget Office. 2015. "The Effects of Potential Cuts in SNAP Spending on Households With Different Amounts of Income." Accessed April 13, 2015. <https://www.cbo.gov/publication/49978>.

Falk, Gene. Congressional Research Service. "The Temporary Assistance for Needy Families (TANF) Block Grant: Responses to Frequently Asked Questions." Last modified October 17, 2013. <http://www.fas.org/sgp/crs/misc/RL32760.pdf>.

Library of Congress. "Congressional Research Service." <http://www.loc.gov/crsinfo/about/>.

Office of Management and Budget. "Fiscal Year 2013 Historical Tables: Budget of the U.S. Government." <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/hist.pdf>.

Tax Policy Center: Urban Institute and Brookings Institution. "The Tax Policy Briefing Book: Taxation and the Family: What is the Earned Income Tax Credit?" <http://www.taxpolicycenter.org/briefing-book/key-elements/family/eitc.cfm>.

Glossary

earned income tax credit (EITC)

a method of assisting the working poor through the tax system

Medicaid

a federal–state joint program enacted in 1965 that provides medical insurance for certain (not all) low-income people, including the near-poor as well as those below the poverty line, and focusing on low-income families with children, the low-income elderly, and the disabled

near-poor

those who have incomes just above the poverty line

safety net

the group of government programs that provide assistance to the poor and the near-poor

Supplemental Nutrition Assistance Program (SNAP)

a federally funded program, started in 1964, in which each month poor people receive SNAP cards they can use to buy food

Income Inequality: Measurement and Causes

By the end of this section, you will be able to:

- Explain the distribution of income, and analyze the sources of income inequality in a market economy
- Measure income distribution in quintiles
- Calculate and graph a Lorenz curve
- Show income inequality through demand and supply diagrams

Poverty levels can be subjective based on the overall income levels of a country; typically poverty is measured based on a percentage of the median income. Income inequality, however, has to do with the distribution of that income, in terms of which group receives the most or the least income. Income inequality involves comparing those with high incomes, middle incomes, and low incomes—not just looking at those below or near the poverty line. In turn, measuring income inequality means dividing up the population into various groups and then comparing the groups, a task that can be carried out in several ways, as the next Clear It Up feature shows.

Note:

How do you separate poverty and income inequality?

Poverty can change even when inequality does not move at all. Imagine a situation in which income for everyone in the population declines by 10%. Poverty would rise, since a greater share of the population would now fall below the poverty line. However, inequality would be the same, because everyone suffered the same proportional loss. Conversely, a general rise in income levels over time would keep inequality the same, but reduce poverty.

It is also possible for income inequality to change without affecting the poverty rate. Imagine a situation in which a large number of people who already have high incomes increase their incomes by even more. Inequality would rise as a result—but the number of people below the poverty line would remain unchanged.

Why did inequality of household income increase in the United States in recent decades? Indeed, a trend toward greater income inequality has occurred in many countries around the world, although the effect has been more powerful in the U.S. economy. Economists have focused their explanations for the increasing inequality on two factors that changed more or less continually from the 1970s into the 2000s. One set of explanations focuses on the changing shape of American households; the other focuses on greater inequality of wages, what some economists call “winner take all” labor markets. We will begin with how we measure inequality, and then consider the explanations for growing inequality in the United States.

Measuring Income Distribution by Quintiles

One common way of measuring income inequality is to rank all households by income, from lowest to highest, and then to divide all households into five groups with equal numbers of people, known as **quintiles**. This calculation allows for measuring the distribution of income among the five groups compared to the total. The first quintile is the lowest fifth or 20%, the second quintile

is the next lowest, and so on. Income inequality can be measured by comparing what share of the total income is earned by each quintile.

U.S. income distribution by quintile appears in [\[link\]](#). In 2011, for example, the bottom quintile of the income distribution received 3.2% of income; the second quintile received 8.4%; the third quintile, 14.3%; the fourth quintile, 23.0%; and the top quintile, 51.14%. The final column of [\[link\]](#) shows what share of income went to households in the top 5% of the income distribution: 22.3% in 2011. Over time, from the late 1960s to the early 1980s, the top fifth of the income distribution typically received between about 43% to 44% of all income. The share of income that the top fifth received then begins to rise. According to the Census Bureau, much of this increase in the share of income going to the top fifth can be traced to an increase in the share of income going to the top 5%. The quintile measure shows how income inequality has increased in recent decades.

Year	Lowest Quintile	Second Quintile	Third Quintile	Fourth Quintile	Highest Quintile	Top 5%
1967	4.0	10.8	17.3	24.2	43.6	17.2
1970	4.1	10.8	17.4	24.5	43.3	16.6
1975	4.3	10.4	17.0	24.7	43.6	16.5
1980	4.2	10.2	16.8	24.7	44.1	16.5
1985	3.9	9.8	16.2	24.4	45.6	17.6
1990	3.8	9.6	15.9	24.0	46.6	18.5
1995	3.7	9.1	15.2	23.3	48.7	21.0
2000	3.6	8.9	14.8	23.0	49.8	22.1
2005	3.4	8.6	14.6	23.0	50.4	22.2
2010	3.3	8.5	14.6	23.4	50.3	21.3
2013	3.2	8.4	14.4	23.0	51	22.2

Share of Aggregate Income Received by Each Fifth and Top 5% of Households, 1967–2013(Source: U.S. Census Bureau, Table 2)

It can also be useful to divide the income distribution in ways other than quintiles; for example, into tenths or even into percentiles (that is, hundredths). A more detailed breakdown can provide

additional insights. For example, the last column of [\[link\]](#) shows the income received by the top 5% percent of the income distribution. Between 1980 and 2013, the share of income going to the top 5% increased by 5.7 percentage points (from 16.5% in 1980 to 22.2% in 2013). From 1980 to 2013 the share of income going to the top quintile increased by 7.0 percentage points (from 44.1% in 1980 to 51% in 2013). Thus, the top 20% of householders (the fifth quintile) received over half (51%) of all the income in the United States in 2013.

Lorenz Curve

The data on income inequality can be presented in various ways. For example, you could draw a bar graph that showed the share of income going to each fifth of the income distribution. [\[link\]](#) presents an alternative way of showing inequality data in what is called a **Lorenz curve**. The Lorenz curve shows the cumulative share of population on the horizontal axis and the cumulative percentage of total income received on the vertical axis.

The Lorenz Curve



A Lorenz curve graphs the cumulative shares of income received by everyone up to a certain quintile. The income distribution in 1980 was closer to the perfect equality line than the income distribution in 2011—that is, the U.S. income distribution became more unequal over time.

Every Lorenz curve diagram begins with a line sloping up at a 45-degree angle, shown as a dashed line in [\[link\]](#). The points along this line show what perfect equality of the income distribution looks like. It would mean, for example, that the bottom 20% of the income distribution receives 20% of the total income, the bottom 40% gets 40% of total income, and so on. The other lines reflect actual U.S. data on inequality for 1980 and 2011.

The trick in graphing a Lorenz curve is that you must change the shares of income for each specific quintile, which are shown in the first column of numbers in [\[link\]](#), into cumulative income, shown in the second column of numbers. For example, the bottom 40% of the cumulative

income distribution will be the sum of the first and second quintiles; the bottom 60% of the cumulative income distribution will be the sum of the first, second, and third quintiles, and so on. The final entry in the cumulative income column needs to be 100%, because by definition, 100% of the population receives 100% of the income.

Income Category	Share of Income in 1980 (%)	Cumulative Share of Income in 1980 (%)	Share of Income in 2013 (%)	Cumulative Share of Income in 2013 (%)
First quintile	4.2	4.2	3.2	3.2
Second quintile	10.2	14.4	8.4	11.6
Third quintile	16.8	31.2	14.4	26.0
Fourth quintile	24.7	55.9	23.0	49.0
Fifth quintile	44.1	100.0	51.0	100.0

Calculating the Lorenz Curve

In a Lorenz curve diagram, a more unequal distribution of income will loop farther down and away from the 45-degree line, while a more equal distribution of income will move the line closer to the 45-degree line. The greater inequality of the U.S. income distribution between 1980 and 2013 is illustrated in [\[link\]](#) because the Lorenz curve for 2013 is farther from the 45-degree line than the Lorenz curve for 1980. The Lorenz curve is a useful way of presenting the quintile data that provides an image of all the quintile data at once. The next Clear It Up feature shows how income inequality differs in various countries compared to the United States.

Note:

How does economic inequality vary around the world?

The U.S. economy has a relatively high degree of income inequality by global standards. As [\[link\]](#) shows, based on a variety of national surveys done for a selection of years in the last five years of the 2000s (with the exception of Germany, and adjusted to make the measures more comparable), the U.S. economy has greater inequality than Germany (along with most Western European countries). The region of the world with the highest level of income inequality is Latin America, illustrated in the numbers for Brazil and Mexico. The level of inequality in the United

States is lower than in some of the low-income countries of the world, like China and Nigeria, or some middle-income countries like the Russian Federation. However, not all poor countries have highly unequal income distributions; India provides a counterexample.

Country	Survey Year	First Quintile	Second Quintile	Third Quintile	Fourth Quintile	Fifth Quintile
United States	2013	3.2%	8.4%	14.4%	23.0%	51.0%
Germany	2000	8.5%	13.7%	17.8%	23.1%	36.9%
Brazil	2009	2.9%	7.1%	12.4%	19.0%	58.6%
Mexico	2010	4.9%	8.8%	13.3%	20.2%	52.8%
China	2009	4.7%	9.7%	15.3%	23.2%	47.1%
India	2010	8.5%	12.1%	15.7%	20.8%	42.8%
Russia	2009	6.1%	10.4%	14.8%	21.3%	47.1%
Nigeria	2010	4.4%	8.3%	13.0%	20.3%	54.0%

Income Distribution in Select Countries(Source: U.S. data from U.S. Census Bureau Table 2. Other data from The World Bank Poverty and Inequality Data Base, <http://databank.worldbank.org/data/views/reports/tableview.aspx#>)

Note:

Visit this [website](#) to watch a video of wealth inequality across the world.



Causes of Growing Inequality: The Changing Composition of American Households

In 1970, 41% of married women were in the labor force, but by 2015, according to the Bureau of Labor Statistics, 56.7% of married women were in the labor force. One result of this trend is that more households have two earners. Moreover, it has become more common for one high earner to marry another high earner. A few decades ago, the common pattern featured a man with relatively high earnings, such as an executive or a doctor, marrying a woman who did not earn as much, like a secretary or a nurse. Often, the woman would leave paid employment, at least for a few years, to raise a family. However, now doctors are marrying doctors and executives are marrying executives, and mothers with high-powered careers are often returning to work while their children are quite young. This pattern of households with two high earners tends to increase the proportion of high-earning households.

According to data in the National Journal, even as two-earner couples have increased, so have single-parent households. Of all U.S. families, 13.1% were headed by single mothers; the poverty rate among single-parent households tends to be relatively high.

These changes in family structure, including the growth of single-parent families who tend to be at the lower end of the income distribution, and the growth of two-career high-earner couples near the top end of the income distribution, account for roughly half of the rise in income inequality across households in recent decades.

Note:

Visit this [website](#) to watch a video that illustrates the distribution of wealth in the United States.



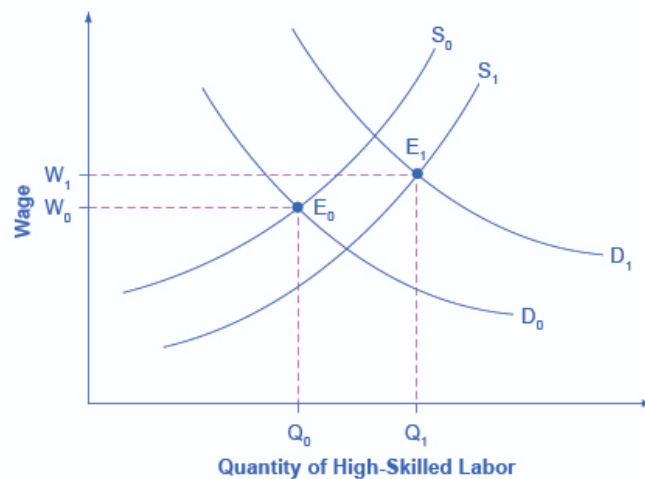
Causes of Growing Inequality: A Shift in the Distribution of Wages

Another factor behind the rise in U.S. income inequality is that earnings have become less equal since the late 1970s. In particular, the earnings of high-skilled labor relative to low-skilled labor have increased. Winner-take-all labor markets result from changes in technology, which have increased global demand for “stars,”—whether the best CEO, doctor, basketball player, or actor. This global demand pushes salaries far above productivity differences versus educational differences. One way to measure this change is to take the earnings of workers with at least a four-year college bachelor’s degree (including those who went on and completed an advanced degree) and divide them by the earnings of workers with only a high school degree. The result is that those in the 25–34 age bracket with college degrees earned about 1.67 times as much as high school

graduates in 2010, up from 1.59 times in 1995, according to U.S. Census data. Winner-take-all labor market theory argues that the salary gap between the median and the top 1 percent is not due to educational differences.

Economists use the demand and supply model to reason through the most likely causes of this shift. According to the National Center for Education Statistics, in recent decades, the supply of U.S. workers with college degrees has increased substantially; for example, 840,000 four-year bachelor's degrees were conferred on Americans in 1970; in 2009–2010, 1,602,480 such degrees were conferred—an increase of about 90%. In [\[link\]](#), this shift in supply to the right, from S_0 to S_1 , should result in a lower equilibrium wage for high-skilled labor. Thus, the increase in the price of high-skilled labor must be explained by a greater demand, like the movement from D_0 to D_1 . Evidently, combining both the increase in supply and in demand has resulted in a shift from E_0 to E_1 , and a resulting higher wage.

Why Would Wages Rise for High-Skilled Labor?



The proportion of workers attending college has increased in recent decades, so the supply curve for high-skilled labor has shifted to the right, from S_0 to S_1 . If the demand for high-skilled labor had remained at D_0 , then this shift in supply would have led to lower wages for high-skilled labor. However, the wages for high-skilled labor, especially if there is a large global demand, have increased even with the shift in supply to the right. The explanation must lie in a shift to the right in demand for high-skilled labor, from D_0 to D_1 . The figure shows how a combination of the shift in supply, from S_0 to S_1 , and the shift in demand, from D_0 to D_1 , led to both an increase in the quantity of high-skilled labor hired and also to a rise in the wage for such labor, from W_0 to W_1 .

What factors would cause the demand for high-skilled labor to rise? The most plausible explanation is that while the explosion in new information and communications technologies over the last several decades has helped many workers to become more productive, the benefits have been especially great for high-skilled workers like top business managers, consultants, and design

professionals. The new technologies have also helped to encourage globalization, the remarkable increase in international trade over the last few decades, by making it more possible to learn about and coordinate economic interactions all around the world. In turn, the rising impact of foreign trade in the U.S. economy has opened up greater opportunities for high-skilled workers to sell their services around the world. And lower-skilled workers have to compete with a larger supply of similarly skilled workers around the globe.

The market for high-skilled labor can be viewed as a race between forces of supply and demand. Additional education and on-the-job training will tend to increase the supply of high-skilled labor and to hold down its relative wage. Conversely, new technology and other economic trends like globalization tend to increase the demand for high-skilled labor and push up its relative wage. The greater inequality of wages can be viewed as a sign that demand for skilled labor is increasing faster than supply. On the other hand, if the supply of lower skilled workers exceeds the demand, then average wages in the lower quintiles of the income distribution will decrease. The combination of forces in the high-skilled and low-skilled labor markets leads to increased income disparity.

Key Concepts and Summary

Measuring inequality involves making comparisons across the entire distribution of income, not just the poor. One way of doing this is to divide the population into groups, like quintiles, and then calculate what share of income is received by each group. An alternative approach is to draw Lorenz curves, which compare the cumulative income actually received to a perfectly equal distribution of income. Income inequality in the United States increased substantially from the late 1970s and early 1980s into the 2000s. The two most common explanations cited by economists are changes in the structure of households that have led to more two-earner couples and single-parent families, and the effect of new information and communications technology on wages.

Self-Check Questions

Exercise:

Problem:

A group of 10 people have the following annual incomes: \$24,000, \$18,000, \$50,000, \$100,000, \$12,000, \$36,000, \$80,000, \$10,000, \$24,000, \$16,000. Calculate the share of total income received by each quintile of this income distribution. Do the top and bottom quintiles in this distribution have a greater or larger share of total income than the top and bottom quintiles of the U.S. income distribution?

Solution:

A useful first step is to rank the households by income, from lowest to highest. Then, since there are 10 households total, the bottom quintile will be the bottom two households, the second quintile will be the third and fourth households, and so on up to the top quintile. The quintiles and percentage of total income for the data provided are shown in the following table. Comparing this distribution to the U.S. income distribution for 2005, the top quintile in the example has a smaller share of total income than in the U.S. distribution and the bottom

quintile has a larger share. This pattern usually means that the income distribution in the example is more equal than the U.S. distribution.

Income	Quintile	% of Total Income
\$10,000	Total first quintile income: \$22,000	6.0%
\$12,000		
\$16,000	Total second quintile income: \$34,000	9.2%
\$18,000		
\$24,000	Total third quintile income: \$48,000	13.0%
\$24,000		
\$36,000	Total fourth quintile income: \$86,000	23.2%
\$50,000		
\$80,000	Total top quintile income: \$180,000	48.6%
\$100,000		
\$370,000	Total Income	

Exercise:

Problem:

[\[link\]](#) shows the share of income going to each quintile of the income distribution for the United Kingdom in 1979 and 1991. Use this data to calculate what the points on a Lorenz curve would be, and sketch the Lorenz curve. How did inequality in the United Kingdom shift over this time period? How can you see the patterns in the quintiles in the Lorenz curves?

Share of Income	1979	1991
-----------------	------	------

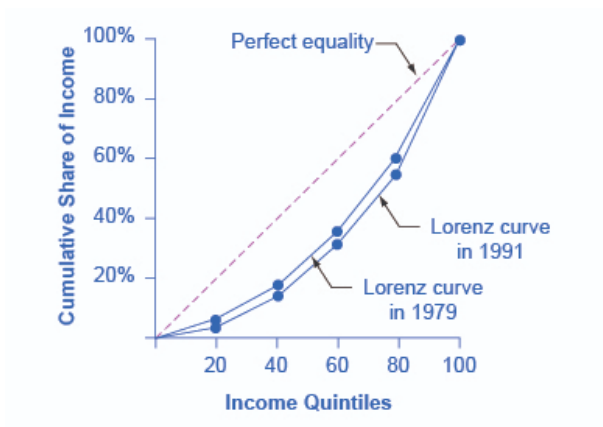
Share of Income	1979	1991
Top quintile	39.7%	42.9%
Fourth quintile	24.8%	22.7%
Middle quintile	17.0%	16.3%
Second quintile	11.5%	11.5%
Bottom quintile	7.0%	6.6%

Income Distribution in the United Kingdom, 1979 and 1991

Solution:

Just from glancing at the quintile information, it is fairly obvious that income inequality increased in the United Kingdom over this time: The top quintile is getting a lot more, and the lowest quintile is getting a bit less. Converting this information into a Lorenz curve, however, is a little trickier, because the Lorenz curve graphs the cumulative distribution, not the amount received by individual quintiles. Thus, as explained in the text, you have to add up the individual quintile data to convert the data to this form. The following table shows the actual calculations for the share of income in 1979 versus 1991. The figure following the table shows the perfect equality line and the Lorenz curves for 1979 and 1991. As shown, the income distribution in 1979 was closer to the perfect equality line than the income distribution in 1991—that is, the United Kingdom income distribution became more unequal over time.

Share of income received	1979	1991
Bottom 20%	7.0%	6.6%
Bottom 40%	18.5%	18.1%
Bottom 60%	35.5%	34.4%
Bottom 80%	60.3%	57.1%
All 100%	100.0%	100.0%



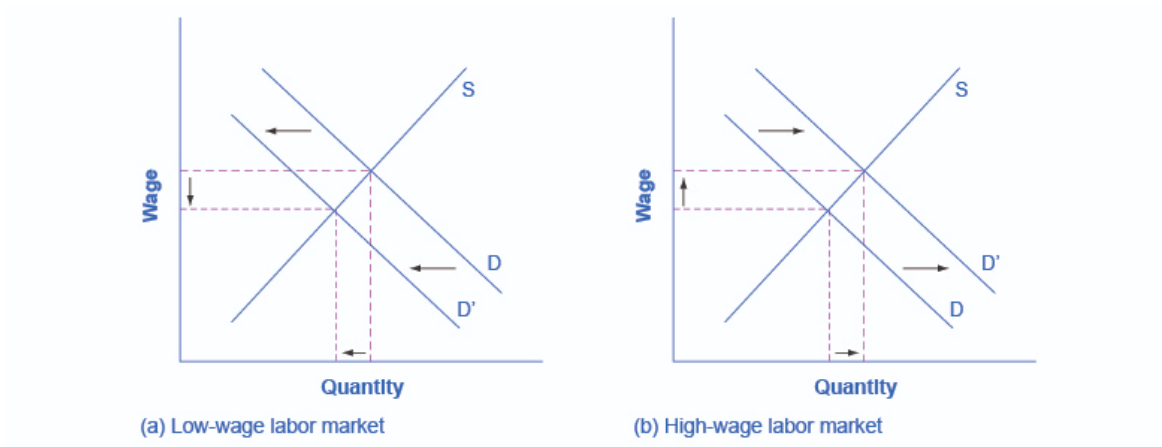
Exercise:

Problem:

Using two demand and supply diagrams, one for the low-wage labor market and one for the high-wage labor market, explain how information technology can increase income inequality if it is a complement to high-income workers like salespeople and managers, but a substitute for low-income workers like file clerks and telephone receptionists.

Solution:

In the market for low-wage labor, information technology shifts the demand for low-wage labor to the left. One reason is that technology can often substitute for low-wage labor in certain kinds of telephone or bookkeeping jobs. In addition, information technology makes it easier for companies to manage connections with low-wage workers in other countries, thus reducing the demand for low-wage workers in the United States. In the market for high-wage labor, information technology shifts the demand for high-wage labor to the right. By using the new information and communications technologies, high-wage labor can become more productive and can oversee more tasks than before. The following figure illustrates these two labor markets. The combination of lower wages for low-wage labor and higher wages for high-wage labor means greater inequality.



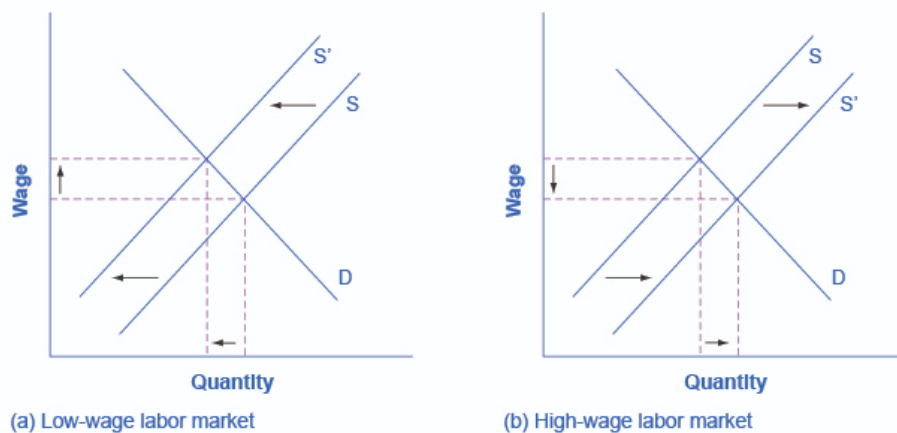
Exercise:

Problem:

Using two demand and supply diagrams, one for the low-wage labor market and one for the high-wage labor market, explain how a program that increased educational levels for a substantial number of low-skill workers could reduce income inequality.

Solution:

In the market for low-wage labor, a skills program will shift supply to the left, which will tend to drive up wages for the remaining low-skill workers. In the market for high-wage labor, a skills program will shift supply to the right (because after the training program there are now more high-skilled workers at every wage), which will tend to drive down wages for high-skill workers. The combination of these two programs will result in a lesser degree of inequality. The following figure illustrates these two labor markets. In the market for high-wage labor, a skills program will shift supply to the right, which will tend to drive down wages for high-skill workers.



Review Questions

Exercise:

Problem: Who is included in the top income quintile?

Exercise:

Problem: What is measured on the two axes of a Lorenz curve?

Exercise:

Problem: If a country had perfect income equality what would the Lorenz curve look like?

Exercise:

Problem:

How has the inequality of income changed in the U.S. economy since the late 1970s?

Exercise:

Problem:

What are some reasons why a certain degree of inequality of income would be expected in a market economy?

Exercise:

Problem:

What are the main reasons economists give for the increase in inequality of incomes?

Critical Thinking Questions

Exercise:

Problem:

Explain how a country may experience greater equality in the distribution of income, yet still experience high rates of poverty *Hint:* Look at the [Clear It Up](#) "How is poverty measured in low-income countries?" and compare to [\[link\]](#).

Exercise:

Problem:

The demand for skilled workers in the United States has been increasing. To increase the supply of skilled workers, many argue that immigration reform to allow more skilled labor into the United States is needed. Explain whether you agree or disagree.

Exercise:**Problem:**

Explain a situation using the supply and demand for skilled labor in which the increased number of college graduates leads to depressed wages. Given the rising cost of going to college, explain why a college education will or will not increase income inequality.

Problems**Exercise:****Problem:**

A group of 10 people have the following annual incomes: \$55,000, \$30,000, \$15,000, \$20,000, \$35,000, \$80,000, \$40,000, \$45,000, \$30,000, \$50,000. Calculate the share of total income received by each quintile of this income distribution. Do the top and bottom quintiles in this distribution have a greater or larger share of total income than the top and bottom quintiles of the U.S. income distribution for 2005?

References

Frank, Robert H., and Philip J. Cook. *The Winner-Take-All Society*. New York: Martin Kessler Books at The Free Press, 1995.

Institute of Education Sciences: National Center for Education Statistics. "Fast Facts: Degrees Conferred by Sex and Race." <http://nces.ed.gov/fastfacts/display.asp?id=72>.

Nhan, Doris. "Census: More in U.S. Report Nontraditional Households." *National Journal*. Last modified May 1, 2012. <http://www.nationaljournal.com/thenextamerica/demographics/census-more-in-u-s-report-nontraditional-households-20120430>.

U.S. Bureau of Labor Statistics: BLS Reports. "Report 1040: Women in the Labor Force: A Databook." Last modified March 26, 2013. <http://www.bls.gov/cps/wlf-databook-2012.pdf>.

U.S. Department of Commerce: United States Census Bureau. "Income: Table H-2. Share of Aggregate Income Received by Each Fifth and Top 5 Percent of Households." <http://www.census.gov/hhes/www/income/data/historical/household/>.

United States Census Bureau. 2014. "2013 Highlights." Accessed April 13, 2015. <http://www.census.gov/hhes/www/poverty/about/overview/>.

United States Census Bureau. 2014. "Historical Income Tables: Households: Table H-2 Share of Aggregate Income Received by Each Fifth and Top 5% of Income. All Races." Accessed April 13, 2015. <http://www.census.gov/hhes/www/income/data/historical/household/>.

Glossary

Lorenz curve

a graph that compares the cumulative income actually received to a perfectly equal distribution of income; it shows the share of population on the horizontal axis and the cumulative percentage of total income received on the vertical axis

quintile

dividing a group into fifths, a method often used to look at distribution of income

Government Policies to Reduce Income Inequality

By the end of this section, you will be able to:

- Explain the arguments for and against government intervention in a market economy
- Identify beneficial ways to reduce the economic inequality in a society
- Show the tradeoff between incentives and income equality

No society should expect or desire complete equality of income at a given point in time, for a number of reasons. First, most workers receive relatively low earnings in their first few jobs, higher earnings as they reach middle age, and then lower earnings after retirement. Thus, a society with people of varying ages will have a certain amount of income inequality. Second, people's preferences and desires differ. Some are willing to work long hours to have income for large houses, fast cars and computers, luxury vacations, and the ability to support children and grandchildren.

These factors all imply that a snapshot of inequality in a given year does not provide an accurate picture of how people's incomes rise and fall over time. Even if some degree of economic inequality is expected at any point in time, how much inequality should there be? There is also the difference between income and wealth, as the following Clear It Up feature explains.

Note:

How do you measure wealth versus income inequality?

Income is a flow of money received, often measured on a monthly or an annual basis; **wealth** is the sum of the value of all assets, including money in bank accounts, financial investments, a pension fund, and the value of a home. In calculating wealth all debts must be subtracted, such as debt owed on a home mortgage and on credit cards. A retired person, for example, may have relatively little income in a given year, other than a pension or Social Security. However, if that person has saved and invested over time, the person's accumulated wealth can be quite substantial. In the United States, the wealth distribution is more unequal than the income distribution, because differences in income can accumulate over time to make even larger differences in wealth. However, the degree of

inequality in the wealth distribution can be measured with the same tools we use to measure the inequality in the income distribution, like quintile measurements. Data on wealth are collected once every three years in the Survey of Consumer Finance.

Even if they cannot answer the question of how much inequality is too much, economists can still play an important role in spelling out policy options and tradeoffs. If a society decides to reduce the level of economic inequality, it has three main sets of tools: redistribution from those with high incomes to those with low incomes; trying to assure that a ladder of opportunity is widely available; and a tax on inheritance.

Redistribution

Redistribution means taking income from those with higher incomes and providing income to those with lower incomes. Earlier in this chapter, we considered some of the key government policies that provide support for the poor: the welfare program TANF, the earned income tax credit, SNAP, and Medicaid. If a reduction in inequality is desired, these programs could receive additional funding.

The programs are paid for through the federal income tax, which is a **progressive tax system** designed in such a way that the rich pay a higher percent in income taxes than the poor. Data from household income tax returns in 2009 shows that the top 1% of households had an average income of \$1,219,700 per year in pre-tax income and paid an average federal tax rate of 28.9%. The **effective income tax**, which is total taxes paid divided by total income (all sources of income such as wages, profits, interest, rental income, and government transfers such as veterans' benefits), was much lower. The effective tax paid by the top 1% of householders was 20.4%, while the bottom two quintiles actually paid negative effective income taxes, because of provisions like the earned income tax credit. News stories occasionally report on a high-income person who has managed to pay very little in taxes, but while such individual cases exist, according to the Congressional Budget Office, the typical pattern is that people with

higher incomes pay a higher average share of their income in federal income taxes.

Of course, the fact that some degree of redistribution occurs now through the federal income tax and government antipoverty programs does not settle the questions of how much redistribution is appropriate, and whether more redistribution should occur.

The Ladder of Opportunity

Economic inequality is perhaps most troubling when it is not the result of effort or talent, but instead is determined by the circumstances under which a child grows up. One child attends a well-run grade school and high school and heads on to college, while parents help out by supporting education and other interests, paying for college, a first car, and a first house, and offering work connections that lead to internships and jobs. Another child attends a poorly run grade school, barely makes it through a low-quality high school, does not go to college, and lacks family and peer support. These two children may be similar in their underlying talents and in the effort they put forth, but their economic outcomes are likely to be quite different.

Public policy can attempt to build a ladder of opportunities so that, even though all children will never come from identical families and attend identical schools, each child has a reasonable opportunity to attain an economic niche in society based on their interests, desires, talents, and efforts. Some of those initiatives include those shown in [\[link\]](#).

Children	College Level	Adults
<ul style="list-style-type: none">• Improved day care	<ul style="list-style-type: none">• Widespread loans and grants for those in financial need	<ul style="list-style-type: none">• Opportunities for retraining and acquiring new skills

Children	College Level	Adults
<ul style="list-style-type: none"> • Enrichment programs for preschoolers 	<ul style="list-style-type: none"> • Public support for a range of institutions from two-year community colleges to large research universities 	<ul style="list-style-type: none"> • Prohibiting discrimination in job markets and housing on the basis of race, gender, age, and disability
<ul style="list-style-type: none"> • Improved public schools 	-	-
<ul style="list-style-type: none"> • After school and community activities 	-	-
<ul style="list-style-type: none"> • Internships and apprenticeships 	-	-

Public Policy Initiatives

The United States has often been called a land of opportunity. Although the general idea of a ladder of opportunity for all citizens continues to exert a powerful attraction, specifics are often quite controversial. Society can experiment with a wide variety of proposals for building a ladder of opportunity, especially for those who otherwise seem likely to start their lives in a disadvantaged position. Such policy experiments need to be carried out in a spirit of open-mindedness, because some will succeed while others will not show positive results or will cost too much to enact on a widespread basis.

Inheritance Taxes

There is always a debate about inheritance taxes. It goes like this: On the one hand, why should people who have worked hard all their lives and

saved up a substantial nest egg not be able to give their money and possessions to their children and grandchildren? In particular, it would seem un-American if children were unable to inherit a family business or a family home. On the other hand, many Americans are far more comfortable with inequality resulting from high-income people who earned their money by starting innovative new companies than they are with inequality resulting from high-income people who have inherited money from rich parents.

The United States does have an **estate tax**—that is, a tax imposed on the value of an inheritance—which suggests a willingness to limit how much wealth can be passed on as an inheritance. However, according to the Center on Budget and Policy Priorities, in 2015 the estate tax applied only to those leaving inheritances of more than \$5.43 million and thus applies to only a tiny percentage of those with high levels of wealth.

The Tradeoff between Incentives and Income Equality

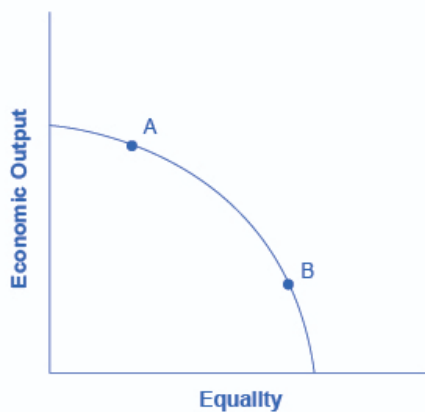
Government policies to reduce poverty or to encourage economic equality, if carried to extremes, can injure incentives for economic output. The poverty trap, for example, defines a situation where guaranteeing a certain level of income can eliminate or reduce the incentive to work. An extremely high degree of redistribution, with very high taxes on the rich, would be likely to discourage work and entrepreneurship. Thus, it is common to draw the tradeoff between economic output and equality, as shown in [\[link\]](#) (a). In this formulation, if society wishes a high level of economic output, like point A, it must also accept a high degree of inequality. Conversely, if society wants a high level of equality, like point B, it must accept a lower level of economic output because of reduced incentives for production.

This view of the tradeoff between economic output and equality may be too pessimistic, and [\[link\]](#) (b) presents an alternate vision. Here, the tradeoff between economic output and equality first slopes up, in the vicinity of choice C, suggesting that certain programs might increase both output and economic equality. For example, the policy of providing free public education has an element of redistribution, since the value of the public schooling received by children of low-income families is clearly higher than what low-income families pay in taxes. A well-educated population,

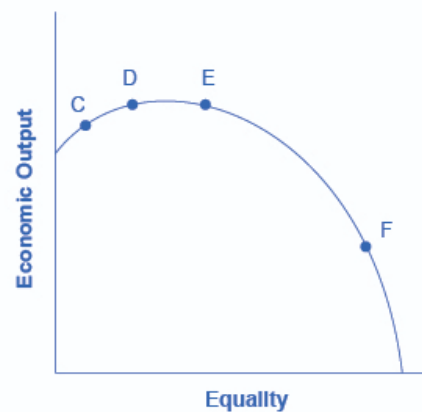
however, is also an enormously powerful factor in providing the skilled workers of tomorrow and helping the economy to grow and expand. In this case, equality and economic growth may complement each other.

Moreover, policies to diminish inequality and soften the hardship of poverty may sustain political support for a market economy. After all, if society does not make some effort toward reducing inequality and poverty, the alternative might be that people would rebel against market forces. Citizens might seek economic security by demanding that their legislators pass laws forbidding employers from ever laying off workers or reducing wages, or laws that would impose price floors and price ceilings and shut off international trade. From this viewpoint, policies to reduce inequality may help economic output by building social support for allowing markets to operate.

The Tradeoff between Incentives and Economic Equality



(a) Greater equality always reduces output



(b) Equality and output rise together but then face a tradeoff

- (a) Society faces a trade-off where any attempt to move toward greater equality, like moving from choice A to B, involves a reduction in economic output. (b) Situations can arise like point C, where it is possible both to increase equality and also to increase economic output, to a choice like D. It may also be possible to increase equality with little impact on economic output, like the movement from choice D to E. However, at some point, too aggressive a push for equality will tend to reduce economic output, as in the shift from E to F.

The tradeoff in [\[link\]](#) (b) then flattens out in the area between points D and E, which reflects the pattern that a number of countries that provide similar levels of income to their citizens—the United States, Canada, the nations of the European Union, Japan, Australia—have different levels of inequality. The pattern suggests that countries in this range could choose a greater or a lesser degree of inequality without much impact on economic output. Only if these countries push for a much higher level of equality, like at point F, will they experience the diminished incentives that lead to lower levels of economic output. In this view, while a danger always exists that an agenda to reduce poverty or inequality can be poorly designed or pushed too far, it is also possible to discover and design policies that improve equality and do not injure incentives for economic output by very much—or even improve such incentives.

Note:**Occupy Wall Street**

The Occupy movement took on a life of its own over the last few months of 2011, bringing to light issues faced by many people on the lower end of the income distribution. The contents of this chapter indicate that there is a significant amount of income inequality in the United States. The question is: What should be done about it?

The Great Recession of 2008–2009 caused unemployment to rise and incomes to fall. Many people attribute the recession to mismanagement of the financial system by bankers and financial managers—those in the 1% of the income distribution—but those in lower quintiles bore the greater burden of the recession through unemployment. This seemed to present the picture of inequality in a different light: the group that seemed responsible for the recession was not the group that seemed to bear the burden of the decline in output. A burden shared can bring a society closer together; a burden pushed off onto others can polarize it.

On one level, the problem with trying to reduce income inequality comes down to whether you still believe in the American Dream. If you believe that one day you will have your American Dream—a large income, large house, happy family, or whatever else you would like to have in life—then you do not necessarily want to prevent anyone else from living out their

dream. You certainly would not want to run the risk that someone would want to take part of your dream away from you. So there is some reluctance to engage in a redistributive policy to reduce inequality. However, when those for whom the likelihood of living the American Dream is very small are considered, there are sound arguments in favor of trying to create greater balance. As the text indicated, a little more income equality, gained through long-term programs like increased education and job training, can increase overall economic output. Then everyone is made better off. And the 1% will not seem like such a small group any more.

Key Concepts and Summary

Policies that can affect the level of economic inequality include redistribution between rich and poor, making it easier for people to climb the ladder of opportunity; and estate taxes, which are taxes on inheritances. Pushing too aggressively for economic equality can run the risk of decreasing economic incentives. However, a moderate push for economic equality can increase economic output, both through methods like improved education and by building a base of political support for market forces.

Self-Check Questions

Exercise:

Problem:

Here is one hypothesis: A well-funded social safety net can increase economic equality but will reduce economic output. Explain why this might be so, and sketch a production possibility curve that shows this tradeoff.

Solution:

A very strong push for economic equality might include extremely high taxes on high-wage earners to pay for extremely large government social payments for the poor. Such a policy could limit

incentives for the high-wage workers, lock the poor into a poverty trap, and thus reduce output. The PPF in this case will have the standard appearance: it will be downward sloping.

Exercise:

Problem:

Here is a second hypothesis: A well-funded social safety net may lead to less regulation of the market economy. Explain why this might be so, and sketch a production possibility curve that shows this tradeoff.

Solution:

For the second hypothesis, a well-funded social safety net might make people feel that even if their company goes bankrupt or they need to change jobs or industries, they will have some degree of protection. As a result, people may be more willing to allow markets to work without interference, and not to lobby as hard for rules that would prevent layoffs, set price controls, or block foreign trade. In this case, safety net programs that increase equality could also allow the market to work more freely in a way that could increase output. In this case, at least some portion of the PPF between equality and economic output would slope up.

Exercise:

Problem:

Which set of policies is more likely to cause a tradeoff between economic output and equality: policies of redistribution or policies aimed at the ladder of opportunity? Explain how the production possibility frontier tradeoff between economic equality and output might look in each case.

Solution:

Pure redistribution is more likely to cause a sharp tradeoff between economic output and equality than policies aimed at the ladder of opportunity. A production possibility frontier showing a strict tradeoff

between economic output and equality will be downward sloping. A PPF showing that it is possible to increase equality, at least to some extent, while either increasing output or at least not diminishing it would have a PPF that first rises, perhaps has a flat area, and then falls.

Exercise:

Problem:

Why is there reluctance on the part of some in the United States to redistribute income so that greater equality can be achieved?

Solution:

Many view the redistribution of income to achieve greater equality as taking away from the rich to pay the poor, or as a “zero sum” game. By taking taxes from one group of people and redistributing them to another, the tax system is robbing some of the American Dream.

Review Questions

Exercise:

Problem:

Identify some public policies that can reduce the level of economic inequality.

Exercise:

Problem:

Describe how a push for economic equality might reduce incentives to work and produce output. Then describe how a push for economic inequality might not have such effects.

Critical Thinking Questions

Exercise:

Problem:

What do you think is more important to focus on when considering inequality: income inequality or wealth inequality?

Exercise:**Problem:**

To reduce income inequality, should the marginal tax rates on the top 1% be increased?

Exercise:**Problem:**

Redistribution of income occurs through the federal income tax and government antipoverty programs. Explain whether or not this level of redistribution is appropriate and whether more redistribution should occur.

Exercise:**Problem:**

How does a society or a country make the decision about the tradeoff between equality and economic output? *Hint:* Think about the political system.

Exercise:**Problem:**

Explain what the long- and short-term consequences are of not promoting equality or working to reduce poverty.

References

Board of Governors of the Federal Reserve System. "Research Resources: Survey of Consumer Finances." Last modified December 13, 2013.
<http://www.federalreserve.gov/econresdata/scf/scfindex.htm>.

Congressional Budget Office. "The Distribution of Household Income and Federal Taxes, 2008 and 2009." Last modified July 10, 2012.
<http://www.cbo.gov/publication/43373>.

Huang, Chye-Ching, and Nathaniel Frentz. "Myths and Realities About the Estate Tax." *Center on Budget and Policy Priorities*. Last modified August 29, 2013. <http://www.cbpp.org/files/estatetaxmyths.pdf>.

Glossary

effective income tax

percentage of total taxes paid divided by total income

estate tax

a tax imposed on the value of an inheritance

income

a flow of money received, often measured on a monthly or an annual basis

progressive tax system

a tax system in which the rich pay a higher percentage of their income in taxes, rather than a higher absolute amount

redistribution

taking income from those with higher incomes and providing income to those with lower incomes

wealth

the sum of the value of all assets, including money in bank accounts, financial investments, a pension fund, and the value of a home

Introduction to International Trade

class="introduction"

Apple or Samsung iPhone?

While the iPhone is readily recognized as an Apple product, 26% of the component costs in it come from components made by rival phone-maker, Samsung. In international trade, there are often “conflicts” like this as each country or company focuses on what it does best.

(Credit: modification of work by Yutaka Tsutano
Creative Commons)

**Note:****Just Whose iPhone Is It?**

The iPhone is a global product. Apple does not manufacture the iPhone components, nor does it assemble them. The assembly is done by Foxconn Corporation, a Taiwanese company, at its factory in Sengzhen, China. But, Samsung, the electronics firm and competitor to Apple, actually supplies many of the parts that make up an iPhone—about 26%. That means, that Samsung is both the biggest supplier and biggest competitor for Apple. Why do these two firms work together to produce the iPhone? To understand the economic logic behind international trade, you have to accept, as these firms do, that trade is about mutually beneficial exchange. Samsung is one of the world's largest electronics parts suppliers. Apple lets Samsung focus on making the best parts, which allows Apple to concentrate on its strength—designing elegant products that are easy to use. If each company (and by extension each country) focuses on what it does best, there will be gains for all through trade.

Note:**Introduction to International Trade**

In this chapter, you will learn about:

- Absolute and Comparative Advantage
- What Happens When a Country Has an Absolute Advantage in All Goods
- Intra-industry Trade between Similar Economies
- The Benefits of Reducing Barriers to International Trade

We live in a global marketplace. The food on your table might include fresh fruit from Chile, cheese from France, and bottled water from Scotland. Your wireless phone might have been made in Taiwan or Korea. The clothes you wear might be designed in Italy and manufactured in China. The toys you give to a child might have come from India. The car you drive might come from Japan, Germany, or Korea. The gasoline in the tank might be refined from crude oil from Saudi Arabia, Mexico, or Nigeria. As a worker, if your job is involved with farming, machinery, airplanes, cars, scientific instruments, or many other technology-related industries, the odds are good that a hearty proportion of the sales of your employer—and hence the money that pays your salary—comes from export sales. We are all linked by international trade, and the volume of that trade has grown dramatically in the last few decades.

The first wave of globalization started in the nineteenth century and lasted up to the beginning of World War I. Over that time, global exports as a share of global GDP rose from less than 1% of GDP in 1820 to 9% of GDP in 1913. As the Nobel Prize-winning economist Paul Krugman of Princeton University wrote in 1995:

"It is a late-twentieth-century conceit that we invented the global economy just yesterday. In fact, world markets achieved an impressive degree of integration during the second half of the nineteenth century. Indeed, if one wants a specific date for the beginning of a truly global economy, one might well choose 1869, the year in which both the Suez Canal and the Union

Pacific railroad were completed. By the eve of the First World War steamships and railroads had created markets for standardized commodities, like wheat and wool, that were fully global in their reach. Even the global flow of information was better than modern observers, focused on electronic technology, tend to realize: the first submarine telegraph cable was laid under the Atlantic in 1858, and by 1900 all of the world's major economic regions could effectively communicate instantaneously."

This first wave of globalization crashed to a halt in the beginning of the twentieth century. World War I severed many economic connections. During the Great Depression of the 1930s, many nations misguidedly tried to fix their own economies by reducing foreign trade with others. World War II further hindered international trade. Global flows of goods and financial capital rebuilt themselves only slowly after World War II. It was not until the early 1980s that global economic forces again became as important, relative to the size of the world economy, as they were before World War I.

Absolute and Comparative Advantage

By the end of this section, you will be able to:

- Define absolute advantage, comparative advantage, and opportunity costs
- Explain the gains of trade created when a country specializes

The American statesman Benjamin Franklin (1706–1790) once wrote: “No nation was ever ruined by trade.” Many economists would express their attitudes toward international trade in an even more positive manner. The evidence that international trade confers overall benefits on economies is pretty strong. Trade has accompanied economic growth in the United States and around the world. Many of the national economies that have shown the most rapid growth in the last few decades—for example, Japan, South Korea, China, and India—have done so by dramatically orienting their economies toward international trade. There is no modern example of a country that has shut itself off from world trade and yet prospered. To understand the benefits of trade, or why we trade in the first place, we need to understand the concepts of comparative and absolute advantage.

In 1817, David Ricardo, a businessman, economist, and member of the British Parliament, wrote a treatise called *On the Principles of Political Economy and Taxation*. In this treatise, Ricardo argued that specialization and free trade benefit all trading partners, even those that may be relatively inefficient. To see what he meant, we must be able to distinguish between absolute and comparative advantage.

A country has an **absolute advantage** in producing a good over another country if it uses fewer resources to produce that good. Absolute advantage can be the result of a country’s natural endowment. For example, extracting oil in Saudi Arabia is pretty much just a matter of “drilling a hole.” Producing oil in other countries can require considerable exploration and costly technologies for drilling and extraction—if indeed they have any oil at all. The United States has some of the richest farmland in the world, making it easier to grow corn and wheat than in many other countries. Guatemala and Colombia have climates especially suited for growing coffee. Chile and Zambia have some of the world’s richest copper mines.

As some have argued, “geography is destiny.” Chile will provide copper and Guatemala will produce coffee, and they will trade. When each country has a product others need and it can be produced with fewer resources in one country over another, then it is easy to imagine all parties benefitting from trade. However, thinking about trade just in terms of geography and absolute advantage is incomplete. Trade really occurs because of comparative advantage.

Recall from the chapter [Choice in a World of Scarcity](#) that a country has a comparative advantage when a good can be produced at a lower cost in terms of other goods. The question each country or company should be asking when it trades is this: “What do we give up to produce this good?” It should be no surprise that the concept of comparative advantage is based on this idea of opportunity cost from [Choice in a World of Scarcity](#). For example, if Zambia focuses its resources on producing copper, its labor, land and financial resources cannot be used to produce other goods such as corn. As a result, Zambia gives up the opportunity to produce corn. How do we quantify the cost in terms of other goods? Simplify the problem and assume that Zambia just needs labor to produce copper and corn. The companies that produce either copper or corn tell you that it takes 10 hours to mine a ton of copper and 20 hours to harvest a bushel of corn. This means the opportunity cost of producing a ton of copper is 2 bushels of corn. The next section develops absolute and comparative advantage in greater detail and relates them to trade.

Note:

Visit this [website](#) for a list of articles and podcasts pertaining to international trade topics.



A Numerical Example of Absolute and Comparative Advantage

Consider a hypothetical world with two countries, Saudi Arabia and the United States, and two products, oil and corn. Further assume that consumers in both countries desire both these goods. These goods are homogeneous, meaning that consumers/producers cannot differentiate between corn or oil from either country. There is only one resource available in both countries, labor hours. Saudi Arabia can produce oil with fewer resources, while the United States can produce corn with fewer resources. [\[link\]](#) illustrates the advantages of the two countries, expressed in terms of how many hours it takes to produce one unit of each good.

Country	Oil (hours per barrel)	Corn (hours per bushel)
Saudi Arabia	1	4
United States	2	1

How Many Hours It Takes to Produce Oil and Corn

In [\[link\]](#), Saudi Arabia has an absolute advantage in the production of oil because it only takes an hour to produce a barrel of oil compared to two hours in the United States. The United States has an absolute advantage in the production of corn.

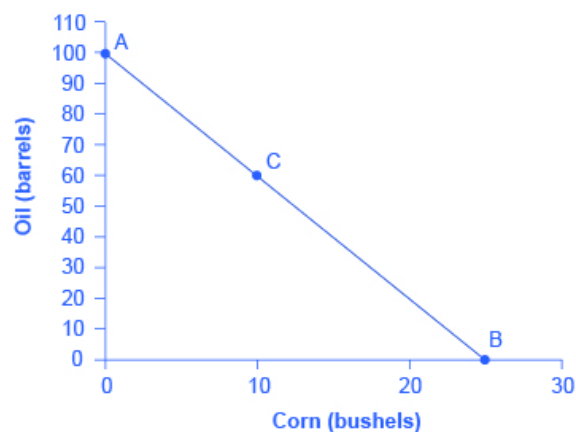
To simplify, let's say that Saudi Arabia and the United States each have 100 worker hours (see [\[link\]](#)). We illustrate what each country is capable of producing on its own using a production possibility frontier (PPF) graph,

shown in [\[link\]](#). Recall from [Choice in a World of Scarcity](#) that the production possibilities frontier shows the maximum amount that each country can produce given its limited resources, in this case workers, and its level of technology.

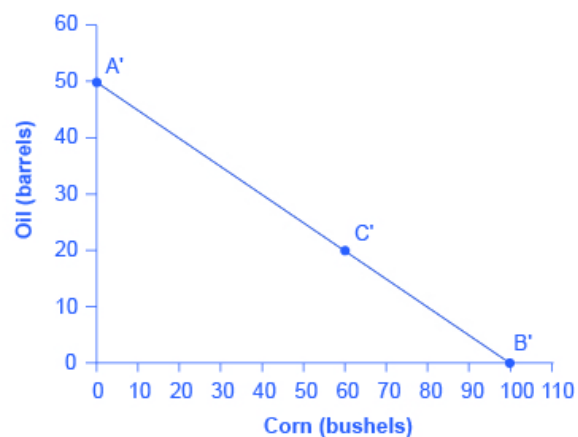
Country	Oil Production using 100 worker hours (barrels)		Corn Production using 100 worker hours (bushels)
Saudi Arabia	100	or	25
United States	50	or	100

Production Possibilities before Trade

Production Possibilities Frontiers



(a) Saudi Arabia



(b) The United States

(a) Saudi Arabia can produce 100 barrels of oil at maximum and zero corn (point A), or 25 bushels of corn and zero oil (point B). It can also produce other combinations of oil and corn if it wants to consume both

goods, such as at point C. Here it chooses to produce/consume 60 barrels of oil, leaving 40 work hours that can be allocated to producing 10 bushels of corn, using the data in [\[link\]](#). (b) If the United States produces only oil, it can produce, at maximum, 50 barrels and zero corn (point A'), or at the other extreme, it can produce a maximum of 100 bushels of corn and no oil (point B'). Other combinations of both oil and corn are possible, such as point C'. All points above the frontiers are impossible to produce given the current level of resources and technology.

Arguably Saudi and U.S. consumers desire both oil and corn to live. Let's say that before trade occurs, both countries produce and consume at point C or C'. Thus, before trade, the Saudi Arabian economy will devote 60 worker hours to produce oil, as shown in [\[link\]](#). Given the information in [\[link\]](#), this choice implies that it produces/consumes 60 barrels of oil. With the remaining 40 worker hours, since it needs four hours to produce a bushel of corn, it can produce only 10 bushels. To be at point C', the U.S. economy devotes 40 worker hours to produce 20 barrels of oil and the remaining worker hours can be allocated to produce 60 bushels of corn.

Country	Oil Production (barrels)	Corn Production (bushels)
Saudi Arabia (C)	60	10
United States (C')	20	60
Total World Production	80	70

Production before Trade

The slope of the production possibility frontier illustrates the opportunity cost of producing oil in terms of corn. Using all its resources, the United States can produce 50 barrels of oil *or* 100 bushels of corn. So the opportunity cost of one barrel of oil is two bushels of corn—or the slope is $1/2$. Thus, in the U.S. production possibility frontier graph, every increase in oil production of one barrel implies a decrease of two bushels of corn. Saudi Arabia can produce 100 barrels of oil *or* 25 bushels of corn. The opportunity cost of producing one barrel of oil is the loss of $1/4$ of a bushel of corn that Saudi workers could otherwise have produced. In terms of corn, notice that Saudi Arabia gives up the least to produce a barrel of oil. These calculations are summarized in [\[link\]](#).

Country	Opportunity cost of one unit — Oil (in terms of corn)	Opportunity cost of one unit — Corn (in terms of oil)
Saudi Arabia	$\frac{1}{4}$	4
United States	2	$\frac{1}{2}$

Opportunity Cost and Comparative Advantage

Again recall that comparative advantage was defined as the opportunity cost of producing goods. Since Saudi Arabia gives up the least to produce a barrel of oil, ($\frac{1}{4} < 2$ in [\[link\]](#)) it has a comparative advantage in oil production. The United States gives up the least to produce a bushel of corn, so it has a comparative advantage in corn production.

In this example, there is symmetry between absolute and comparative advantage. Saudi Arabia needs fewer worker hours to produce oil (absolute advantage, see [\[link\]](#)), and also gives up the least in terms of other goods to produce oil (comparative advantage, see [\[link\]](#)). Such symmetry is not always the case, as we will show after we have discussed gains from trade fully. But first, read the following Clear It Up feature to make sure you understand why the PPF line in the graphs is straight.

Note:

Can a production possibility frontier be straight?

When you first met the production possibility frontier (PPF) in the chapter on [Choice in a World of Scarcity](#), it was drawn with an outward-bending shape. This shape illustrated that as inputs were transferred from producing one good to another—like from education to health services—there were increasing opportunity costs. In the examples in this chapter, the PPFs are drawn as straight lines, which means that opportunity costs are constant. When a marginal unit of labor is transferred away from growing corn and toward producing oil, the decline in the quantity of corn and the increase in the quantity of oil is always the same. In reality this is possible only if the contribution of additional workers to output did not change as the scale of production changed. The linear production possibilities frontier is a less realistic model, but a straight line simplifies calculations. It also illustrates economic themes like absolute and comparative advantage just as clearly.

Gains from Trade

Consider the trading positions of the United States and Saudi Arabia after they have specialized and traded. Before trade, Saudi Arabia produces/consumes 60 barrels of oil and 10 bushels of corn. The United States produces/consumes 20 barrels of oil and 60 bushels of corn. Given their current production levels, if the United States can trade an amount of corn fewer than 60 bushels and receives in exchange an amount of oil greater than 20 barrels, it will **gain from trade**. With trade, the United States can consume more of both goods than it did without specialization

and trade. (Recall that the chapter [Welcome to Economics!](#) defined specialization as it applies to workers and firms. Specialization is also used to describe the occurrence when a country shifts resources to focus on producing a good that offers comparative advantage.) Similarly, if Saudi Arabia can trade an amount of oil less than 60 barrels and receive in exchange an amount of corn greater than 10 bushels, it will have more of both goods than it did before specialization and trade. [\[link\]](#) illustrates the range of trades that would benefit both sides.

The U.S. Economy, after Specialization, Will Benefit If It:	The Saudi Arabian Economy, after Specialization, Will Benefit If It:
Exports no more than 60 bushels of corn	Imports at least 10 bushels of corn
Imports at least 20 barrels of oil	Exports less than 60 barrels of oil

The Range of Trades That Benefit Both the United States and Saudi Arabia

The underlying reason why trade benefits both sides is rooted in the concept of opportunity cost, as the following Clear It Up feature explains. If Saudi Arabia wishes to expand domestic production of corn in a world without international trade, then based on its opportunity costs it must give up four barrels of oil for every one additional bushel of corn. If Saudi Arabia could find a way to give up less than four barrels of oil for an additional bushel of corn (or equivalently, to receive more than one bushel of corn for four barrels of oil), it would be better off.

Note:

What are the opportunity costs and gains from trade?

The range of trades that will benefit each country is based on the country's opportunity cost of producing each good. The United States can produce 100 bushels of corn or 50 barrels of oil. For the United States, the opportunity cost of producing one barrel of oil is two bushels of corn. If we divide the numbers above by 50, we get the same ratio: one barrel of oil is equivalent to two bushels of corn, or $(100/50 = 2$ and $50/50 = 1)$. In a trade with Saudi Arabia, if the United States is going to give up 100 bushels of corn in exports, it must import at least 50 barrels of oil to be just as well off. Clearly, to gain from trade it needs to be able to gain more than a half barrel of oil for its bushel of corn—or why trade at all?

Recall that David Ricardo argued that if each country specializes in its comparative advantage, it will benefit from trade, and total global output will increase. How can we show gains from trade as a result of comparative advantage and specialization? [\[link\]](#) shows the output assuming that each country specializes in its comparative advantage and produces no other good. This is 100% specialization. Specialization leads to an increase in total world production. (Compare the total world production in [\[link\]](#) to that in [\[link\]](#).)

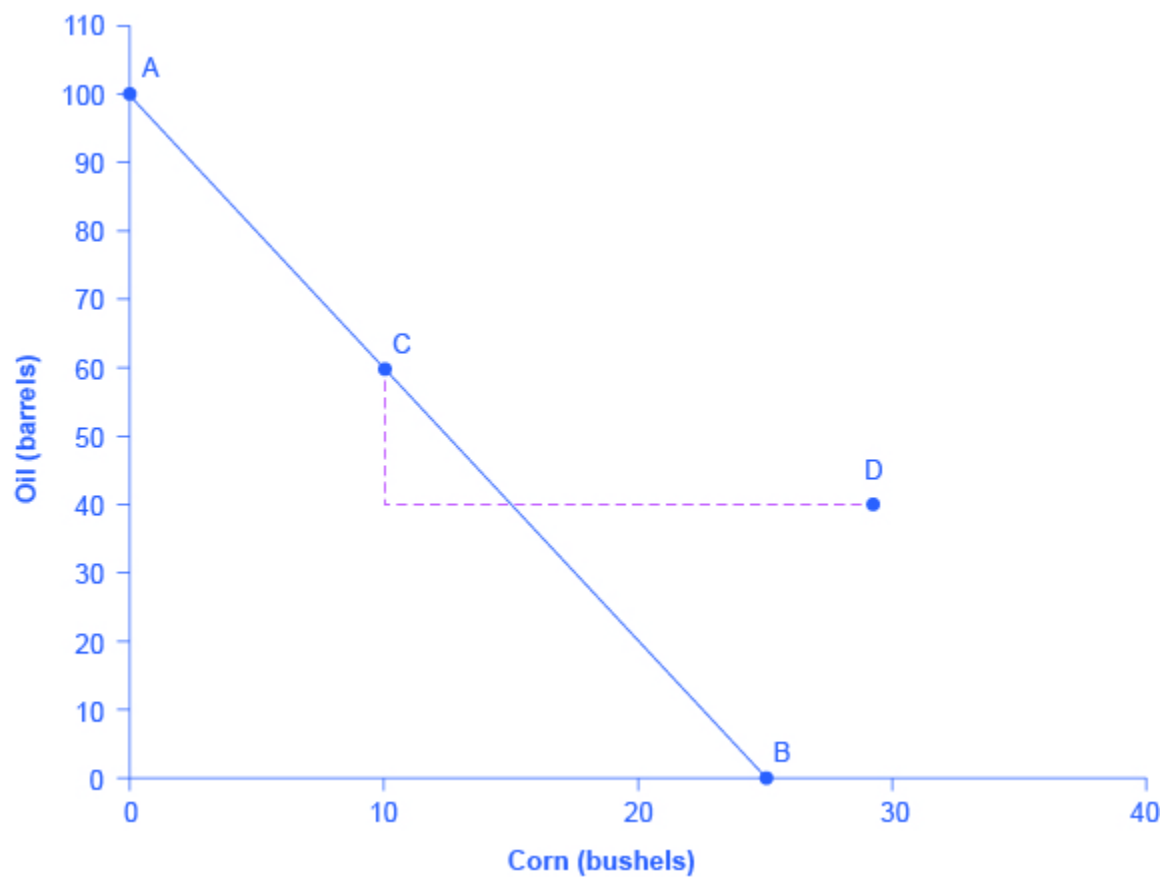
Country	Quantity produced after 100% specialization — Oil (barrels)	Quantity produced after 100% specialization — Corn (bushels)
Saudi Arabia	100	0

Country	Quantity produced after 100% specialization — Oil (barrels)	Quantity produced after 100% specialization — Corn (bushels)
United States	0	100
Total World Production	100	100

How Specialization Expands Output

What if we did not have complete specialization, as in [\[link\]](#)? Would there still be gains from trade? Consider another example, such as when the United States and Saudi Arabia start at C and C', respectively, as shown in [\[link\]](#). Consider what occurs when trade is allowed and the United States exports 20 bushels of corn to Saudi Arabia in exchange for 20 barrels of oil.

Production Possibilities Frontier in Saudi Arabia



Gains from trade of oil can increase only by achieving less from trade of corn. The opposite is true as well: The more gains from trade of corn, the fewer gains from trade of oil.

Starting at point C, which shows Saudi oil production of 60, reduce Saudi Oil domestic oil consumption by 20, since 20 is exported to United States and exchanged for 20 units of corn. This enables Saudi to reach point D, where consumption of oil is now 40 barrels and consumption of corn has increased to 30 (see [\[link\]](#)). Notice that even without 100% specialization, if the “trading price,” in this case 20 barrels of oil for 20 bushels of corn, is greater than the country’s opportunity cost, the Saudis will gain from trade. Since the post-trade consumption point D is beyond its production possibility frontier, hence Saudi Arabia has gained from trade.

Note:

Visit this [website](#) for trade-related data visualizations.



Key Concepts and Summary

A country has an absolute advantage in those products in which it has a productivity edge over other countries; it takes fewer resources to produce a product. A country has a comparative advantage when a good can be produced at a lower cost in terms of other goods. Countries that specialize based on comparative advantage gain from trade.

Self-Check Questions

Exercise:**Problem:**

True or False: The source of comparative advantage must be natural elements like climate and mineral deposits. Explain.

Solution:

False. Anything that leads to different levels of productivity between two economies can be a source of comparative advantage. For example, the education of workers, the knowledge base of engineers and scientists in a country, the part of a split-up value chain where they have their specialized learning, economies of scale, and other factors can all determine comparative advantage.

Exercise:**Problem:**

Brazil can produce 100 pounds of beef or 10 autos; in contrast the United States can produce 40 pounds of beef or 30 autos. Which country has the absolute advantage in beef? Which country has the absolute advantage in producing autos? What is the opportunity cost of producing one pound of beef in Brazil? What is the opportunity cost of producing one pound of beef in the United States?

Solution:

Brazil has the absolute advantage in producing beef and the United States has the absolute advantage in autos. The opportunity cost of producing one pound of beef is $1/10$ of an auto; in the United States it is $3/4$ of an auto.

Exercise:**Problem:**

In France it takes one worker to produce one sweater, and one worker to produce one bottle of wine. In Tunisia it takes two workers to produce one sweater, and three workers to produce one bottle of wine. Who has the absolute advantage in production of sweaters? Who has the absolute advantage in the production of wine? How can you tell?

Solution:

In answering questions like these, it is often helpful to begin by organizing the information in a table, such as in the following table. Notice that, in this case, the productivity of the countries is expressed in terms of how many workers it takes to produce a unit of a product.

Country	One Sweater	One Bottle of wine
France	1 worker	1 worker
Tunisia	2 workers	3 workers

In this example, France has an absolute advantage in the production of both sweaters and wine. You can tell because it takes France less labor to produce a unit of the good.

Review Questions

Exercise:

Problem:

What is absolute advantage? What is comparative advantage?

Exercise:

Problem:

Under what conditions does comparative advantage lead to gains from trade?

Exercise:

Problem:

What factors does Paul Krugman identify that supported the expansion of international trade in the 1800s?

Critical Thinking Questions

Exercise:

Problem:

Are differences in geography behind the differences in absolute advantages?

Exercise:**Problem:**

Why does the United States not have an absolute advantage in coffee?

Exercise:**Problem:**

Look at [\[link\]](#). Compute the opportunity costs of producing sweaters and wine in both France and Tunisia. Who has the lowest opportunity cost of producing sweaters and who has the lowest opportunity cost of producing wine? Explain what it means to have a lower opportunity cost.

Problems**Exercise:****Problem:**

France and Tunisia both have Mediterranean climates that are excellent for producing/harvesting green beans and tomatoes. In France it takes two hours for each worker to harvest green beans and two hours to harvest a tomato. Tunisian workers need only one hour to harvest the tomatoes but four hours to harvest green beans. Assume there are only two workers, one in each country, and each works 40 hours a week.

- a. Draw a production possibilities frontier for each country. *Hint:* Remember the production possibility frontier is the maximum that all workers can produce at a unit of time which, in this problem, is a week.

- b. Identify which country has the absolute advantage in green beans and which country has the absolute advantage in tomatoes.
- c. Identify which country has the comparative advantage.
- d. How much would France have to give up in terms of tomatoes to gain from trade? How much would it have to give up in terms of green beans?

References

Krugman, Paul R. *Pop Internationalism*. The MIT Press, Cambridge. 1996.

Krugman, Paul R. "What Do Undergrads Need to Know about Trade?" *American Economic Review* 83, no. 2. 1993. 23-26.

Ricardo, David. *On the Principles of Political Economy and Taxation*. London: John Murray, 1817.

Ricardo, David. "On the Principles of Political Economy and Taxation." *Library of Economics and Liberty*.
<http://www.econlib.org/library/Ricardo/ricP.html>.

Glossary

absolute advantage

when one country can use fewer resources to produce a good compared to another country; when a country is more productive compared to another country

gain from trade

a country that can consume more than it can produce as a result of specialization and trade

What Happens When a Country Has an Absolute Advantage in All Goods

By the end of this section, you will be able to:

- Show the relationship between production costs and comparative advantage
- Identify situations of mutually beneficial trade
- Identify trade benefits by considering opportunity costs

What happens to the possibilities for trade if one country has an absolute advantage in everything? This is typical for high-income countries that often have well-educated workers, technologically advanced equipment, and the most up-to-date production processes. These high-income countries can produce all products with fewer resources than a low-income country. If the high-income country is more productive across the board, will there still be gains from trade? Good students of Ricardo understand that trade is about mutually beneficial exchange. Even when one country has an absolute advantage in all products, trade can still benefit both sides. This is because gains from trade come from specializing in one's comparative advantage.

Production Possibilities and Comparative Advantage

Consider the example of trade between the United States and Mexico described in [\[link\]](#). In this example, it takes four U.S. workers to produce 1,000 pairs of shoes, but it takes five Mexican workers to do so. It takes one U.S. worker to produce 1,000 refrigerators, but it takes four Mexican workers to do so. The United States has an absolute advantage in productivity with regard to both shoes and refrigerators; that is, it takes fewer workers in the United States than in Mexico to produce both a given number of shoes and a given number of refrigerators.

Country	Number of Workers needed to produce 1,000 units — Shoes	Number of Workers needed to produce 1,000 units — Refrigerators
United States	4 workers	1 worker
Mexico	5 workers	4 workers

Resources Needed to Produce Shoes and Refrigerators

Absolute advantage simply compares the productivity of a worker between countries. It answers the question, “How many inputs do I need to produce shoes in Mexico?” Comparative advantage asks this same question slightly differently. Instead of comparing how many workers it takes to produce a good, it asks, “How much am I giving up to produce this good in this country?” Another way of looking at this is that comparative advantage identifies the good for which the producer’s absolute advantage is relatively larger, or where the producer’s absolute productivity disadvantage is relatively smaller. The United States can produce 1,000 shoes with four-fifths as many workers as Mexico (four versus five), but it can produce 1,000 refrigerators with only one-quarter as many workers (one versus four). So, the comparative advantage of the United States, where its absolute productivity advantage is relatively greatest, lies with refrigerators, and Mexico’s comparative advantage, where its absolute productivity disadvantage is least, is in the production of shoes.

Mutually Beneficial Trade with Comparative Advantage

When nations increase production in their area of comparative advantage and trade with each other, both countries can benefit. Again, the production possibility frontier is a useful tool to visualize this benefit.

Consider a situation where the United States and Mexico each have 40 workers. For example, as [\[link\]](#) shows, if the United States divides its labor so that 40 workers are making shoes, then, since it takes four workers in the

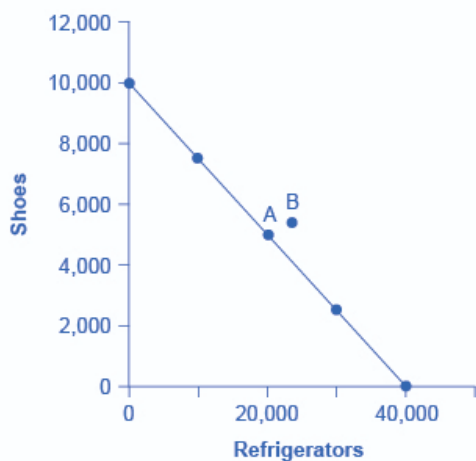
United States to make 1,000 shoes, a total of 10,000 shoes will be produced. (If four workers can make 1,000 shoes, then 40 workers will make 10,000 shoes). If the 40 workers in the United States are making refrigerators, and each worker can produce 1,000 refrigerators, then a total of 40,000 refrigerators will be produced.

Country	Shoe Production — using 40 workers		Refrigerator Production — using 40 workers
United States	10,000 shoes	or	40,000 refrigerators
Mexico	8,000 shoes	or	10,000 refrigerators

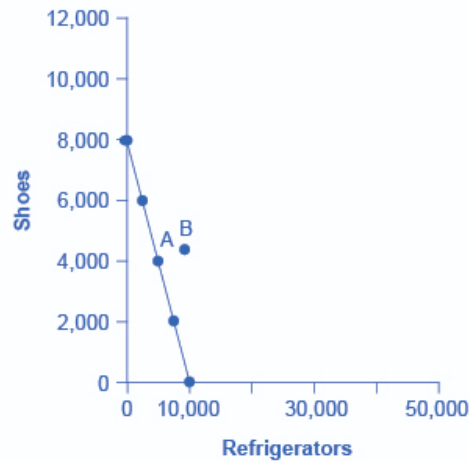
Production Possibilities before Trade with Complete Specialization

As always, the slope of the production possibility frontier for each country is the opportunity cost of one refrigerator in terms of foregone shoe production—when labor is transferred from producing the latter to producing the former (see [\[link\]](#)).

Production Possibility Frontiers



(a) U.S. PPF (40 workers)



(b) Mexico PPF (40 workers)

(a) With 40 workers, the United States can produce either 10,000 shoes and zero refrigerators or 40,000 refrigerators and zero shoes. (b) With 40 workers, Mexico can produce a maximum of 8,000 shoes and zero refrigerators, or 10,000 refrigerators and zero shoes. All other points on the production possibility line are possible combinations of the two goods that can be produced given current resources. Point A on both graphs is where the countries start producing and consuming before trade. Point B is where they end up after trade.

Let's say that, in the situation before trade, each nation prefers to produce a combination of shoes and refrigerators that is shown at point A. [\[link\]](#) shows the output of each good for each country and the total output for the two countries.

Country	Current Shoe Production	Current Refrigerator Production
---------	-------------------------	---------------------------------

Country	Current Shoe Production	Current Refrigerator Production
United States	5,000	20,000
Mexico	4,000	5,000
Total	9,000	25,000

Total Production at Point A before Trade

Continuing with this scenario, each country transfers some amount of labor toward its area of comparative advantage. For example, the United States transfers six workers away from shoes and toward producing refrigerators. As a result, U.S. production of shoes decreases by 1,500 units ($6/4 \times 1,000$), while its production of refrigerators increases by 6,000 (that is, $6/1 \times 1,000$). Mexico also moves production toward its area of comparative advantage, transferring 10 workers away from refrigerators and toward production of shoes. As a result, production of refrigerators in Mexico falls by 2,500 ($10/4 \times 1,000$), but production of shoes increases by 2,000 pairs ($10/5 \times 1,000$). Notice that when both countries shift production toward each of their comparative advantages (what they are relatively better at), their combined production of both goods rises, as shown in [\[link\]](#). The reduction of shoe production by 1,500 pairs in the United States is more than offset by the gain of 2,000 pairs of shoes in Mexico, while the reduction of 2,500 refrigerators in Mexico is more than offset by the additional 6,000 refrigerators produced in the United States.

Country	Shoe Production	Refrigerator Production
---------	-----------------	-------------------------

Country	Shoe Production	Refrigerator Production
United States	3,500	26,000
Mexico	6,000	2,500
Total	9,500	28,500

Shifting Production Toward Comparative Advantage Raises Total Output

This numerical example illustrates the remarkable insight of comparative advantage: even when one country has an absolute advantage in all goods and another country has an absolute disadvantage in all goods, both countries can still benefit from trade. Even though the United States has an absolute advantage in producing both refrigerators and shoes, it makes economic sense for it to specialize in the good for which it has a comparative advantage. The United States will export refrigerators and in return import shoes.

How Opportunity Cost Sets the Boundaries of Trade

This example shows that both parties can benefit from specializing in their comparative advantages and trading. By using the opportunity costs in this example, it is possible to identify the range of possible trades that would benefit each country.

Mexico started out, before specialization and trade, producing 4,000 pairs of shoes and 5,000 refrigerators (see [\[link\]](#) and [\[link\]](#)). Then, in the numerical example given, Mexico shifted production toward its comparative advantage and produced 6,000 pairs of shoes but only 2,500 refrigerators. Thus, if Mexico can export no more than 2,000 pairs of shoes (giving up 2,000 pairs of shoes) in exchange for imports of at least 2,500 refrigerators (a gain of 2,500 refrigerators), it will be able to consume more of both goods than before trade. Mexico will be unambiguously better off. Conversely, the United States started off, before specialization and trade, producing 5,000 pairs of shoes and 20,000 refrigerators. In the example, it

then shifted production toward its comparative advantage, producing only 3,500 shoes but 26,000 refrigerators. If the United States can export no more than 6,000 refrigerators in exchange for imports of at least 1,500 pairs of shoes, it will be able to consume more of both goods and will be unambiguously better off.

The range of trades that can benefit both nations is shown in [\[link\]](#). For example, a trade where the U.S. exports 4,000 refrigerators to Mexico in exchange for 1,800 pairs of shoes would benefit both sides, in the sense that both countries would be able to consume more of both goods than in a world without trade.

The U.S. economy, after specialization, will benefit if it:	The Mexican economy, after specialization, will benefit if it:
<i>Exports</i> fewer than 6,000 refrigerators	<i>Imports</i> at least 2,500 refrigerators
<i>Imports</i> at least 1,500 pairs of shoes	<i>Exports</i> no more than 2,000 pairs of shoes

The Range of Trades That Benefit Both the United States and Mexico

Trade allows each country to take advantage of lower opportunity costs in the other country. If Mexico wants to produce more refrigerators without trade, it must face its domestic opportunity costs and reduce shoe production. If Mexico, instead, produces more shoes and then trades for refrigerators made in the United States, where the opportunity cost of producing refrigerators is lower, Mexico can in effect take advantage of the lower opportunity cost of refrigerators in the United States. Conversely, when the United States specializes in its comparative advantage of refrigerator production and trades for shoes produced in Mexico,

international trade allows the United States to take advantage of the lower opportunity cost of shoe production in Mexico.

The theory of comparative advantage explains why countries trade: they have different comparative advantages. It shows that the gains from international trade result from pursuing comparative advantage and producing at a lower opportunity cost. The following Work It Out feature shows how to calculate absolute and comparative advantage and the way to apply them to a country's production.

Note:

Calculating Absolute and Comparative Advantage

In Canada a worker can produce 20 barrels of oil or 40 tons of lumber. In Venezuela, a worker can produce 60 barrels of oil or 30 tons of lumber.

Country	Oil (barrels)		Lumber (tons)
Canada	20	or	40
Venezuela	60	or	30

- Who has the absolute advantage in the production of oil or lumber? How can you tell?
- Which country has a comparative advantage in the production of oil?
- Which country has a comparative advantage in producing lumber?
- In this example, is absolute advantage the same as comparative advantage, or not?
- In what product should Canada specialize? In what product should Venezuela specialize?

Step 1. Make a table like [\[link\]](#).

Step 2. To calculate absolute advantage, look at the larger of the numbers for each product. One worker in Canada can produce more lumber (40 tons versus 30 tons), so Canada has the absolute advantage in lumber. One worker in Venezuela can produce 60 barrels of oil compared to a worker in Canada who can produce only 20.

Step 3. To calculate comparative advantage, find the opportunity cost of producing one barrel of oil in both countries. The country with the lowest opportunity cost has the comparative advantage. With the same labor time, Canada can produce either 20 barrels of oil or 40 tons of lumber. So in effect, 20 barrels of oil is equivalent to 40 tons of lumber: $20 \text{ oil} = 40 \text{ lumber}$. Divide both sides of the equation by 20 to calculate the opportunity cost of one barrel of oil in Canada. $20/20 \text{ oil} = 40/20 \text{ lumber}$. $1 \text{ oil} = 2 \text{ lumber}$. To produce one additional barrel of oil in Canada has an opportunity cost of 2 lumber. Calculate the same way for Venezuela: $60 \text{ oil} = 30 \text{ lumber}$. Divide both sides of the equation by 60. One oil in Venezuela has an opportunity cost of $1/2 \text{ lumber}$. Because $1/2 \text{ lumber} < 2 \text{ lumber}$, Venezuela has the comparative advantage in producing oil.

Step 4. Calculate the opportunity cost of one lumber by reversing the numbers, with lumber on the left side of the equation. In Canada, 40 lumber is equivalent in labor time to 20 barrels of oil: $40 \text{ lumber} = 20 \text{ oil}$. Divide each side of the equation by 40. The opportunity cost of one lumber is $1/2 \text{ oil}$. In Venezuela, the equivalent labor time will produce 30 lumber or 60 oil: $30 \text{ lumber} = 60 \text{ oil}$. Divide each side by 30. One lumber has an opportunity cost of two oil. Canada has the lower opportunity cost in producing lumber.

Step 5. In this example, absolute advantage is the same as comparative advantage. Canada has the absolute and comparative advantage in lumber; Venezuela has the absolute and comparative advantage in oil.

Step 6. Canada should specialize in what it has a relative lower opportunity cost, which is lumber, and Venezuela should specialize in oil. Canada will be exporting lumber and importing oil, and Venezuela will be exporting oil and importing lumber.

Comparative Advantage Goes Camping

To build an intuitive understanding of how comparative advantage can benefit all parties, set aside examples that involve national economies for a moment and consider the situation of a group of friends who decide to go camping together. The six friends have a wide range of skills and experiences, but one person in particular, Jethro, has done lots of camping before and is also a great athlete. Jethro has an absolute advantage in all aspects of camping: he is faster at carrying a backpack, gathering firewood, paddling a canoe, setting up tents, making a meal, and washing up. So here is the question: Because Jethro has an absolute productivity advantage in everything, should he do all the work?

Of course not! Even if Jethro is willing to work like a mule while everyone else sits around, he, like most mortals, only has 24 hours in a day. If everyone sits around and waits for Jethro to do everything, not only will Jethro be an unhappy camper, but there will not be much output for his group of six friends to consume. The theory of comparative advantage suggests that everyone will benefit if they figure out their areas of comparative advantage—that is, the area of camping where their productivity disadvantage is least, compared to Jethro. For example, it may be that Jethro is 80% faster at building fires and cooking meals than anyone else, but only 20% faster at gathering firewood and 10% faster at setting up tents. In that case, Jethro should focus on building fires and making meals, and others should attend to the other tasks, each according to where their productivity disadvantage is smallest. If the campers coordinate their efforts according to comparative advantage, they can all gain.

Key Concepts and Summary

Even when a country has high levels of productivity in all goods, it can still benefit from trade. Gains from trade come about as a result of comparative advantage. By specializing in a good that it gives up the least to produce, a country can produce more and offer that additional output for sale. If other countries specialize in the area of their comparative advantage as well and trade, the highly productive country is able to benefit from a lower opportunity cost of production in other countries.

Self-Check Question

Exercise:

Problem:

In Germany it takes three workers to make one television and four workers to make one video camera. In Poland it takes six workers to make one television and 12 workers to make one video camera.

- a. Who has the absolute advantage in the production of televisions? Who has the absolute advantage in the production of video cameras? How can you tell?
- b. Calculate the opportunity cost of producing one additional television set in Germany and in Poland. (Your calculation may involve fractions, which is fine.) Which country has a comparative advantage in the production of televisions?
- c. Calculate the opportunity cost of producing one video camera in Germany and in Poland. Which country has a comparative advantage in the production of video cameras?
- d. In this example, is absolute advantage the same as comparative advantage, or not?
- e. In what product should Germany specialize? In what product should Poland specialize?

Solution:

(a) In Germany, it takes fewer workers to make either a television or a video camera. Germany has an absolute advantage in the production of both goods.

(b) Producing an additional television in Germany requires three workers. Shifting those three German workers will reduce video camera production by $\frac{3}{4}$ of a camera. Producing an additional television set in Poland requires six workers, and shifting those workers from the other good reduces output of video cameras by $\frac{6}{12}$ of a camera, or $\frac{1}{2}$. Thus, the opportunity cost of producing televisions is lower in Poland, so Poland has the comparative advantage in the

production of televisions. *Note:* Do not let the fractions like $\frac{3}{4}$ of a camera or $\frac{1}{2}$ of a video camera bother you. If either country was to expand television production by a significant amount—that is, lots more than one unit—then we will be talking about whole cameras and not fractional ones. You can also spot this conclusion by noticing that Poland's absolute disadvantage is relatively lower in televisions, because Poland needs twice as many workers to produce a television but three times as many to produce a video camera, so the product with the relatively lower absolute disadvantage is Poland's comparative advantage.

(c) Producing a video camera in Germany requires four workers, and shifting those four workers away from television production has an opportunity cost of $\frac{4}{3}$ television sets. Producing a video camera in Poland requires 12 workers, and shifting those 12 workers away from television production has an opportunity cost of two television sets. Thus, the opportunity cost of producing video cameras is lower in Germany, and video cameras will be Germany's comparative advantage.

(d) In this example, absolute advantage differs from comparative advantage. Germany has the absolute advantage in the production of both goods, but Poland has a comparative advantage in the production of televisions. (e) Germany should specialize, at least to some extent, in the production of video cameras, export video cameras, and import televisions. Conversely, Poland should specialize, at least to some extent, in the production of televisions, export televisions, and import video cameras.

Review Questions

Exercise:

Problem:

Is it possible to have a comparative advantage in the production of a good but not to have an absolute advantage? Explain.

Exercise:

Problem: How does comparative advantage lead to gains from trade?

Critical Thinking Questions

Exercise:

Problem:

You just overheard your friend say the following: “Poor countries like Malawi have no absolute advantages. They have poor soil, low investments in formal education and hence low-skill workers, no capital, and no natural resources to speak of. Because they have no advantage, they cannot benefit from trade.” How would you respond?

Exercise:

Problem:

Look at [\[link\]](#). Is there a range of trades for which there will be no gains?

Exercise:

Problem:

You just got a job in Washington, D.C. You move into an apartment with some acquaintances. All your roommates, however, are slackers and do not clean up after themselves. You, on the other hand, can clean faster than each of them. You determine that you are 70% faster at dishes and 10% faster with vacuuming. All of these tasks have to be done daily. Which jobs should you assign to your roommates to get the most free time overall? Assume you have the same number of hours to devote to cleaning. Now, since you are faster, you seem to get done quicker than your roommate. What sorts of problems may this create? Can you imagine a trade-related analogy to this problem?

Problems

Exercise:**Problem:**

In Japan, one worker can make 5 tons of rubber or 80 radios. In Malaysia, one worker can make 10 tons of rubber or 40 radios.

- a. Who has the absolute advantage in the production of rubber or radios? How can you tell?
- b. Calculate the opportunity cost of producing 80 additional radios in Japan and in Malaysia. (Your calculation may involve fractions, which is fine.) Which country has a comparative advantage in the production of radios?
- c. Calculate the opportunity cost of producing 10 additional tons of rubber in Japan and in Malaysia. Which country has a comparative advantage in producing rubber?
- d. In this example, does each country have an absolute advantage and a comparative advantage in the same good?
- e. In what product should Japan specialize? In what product should Malaysia specialize?

Exercise:**Problem:**

Review the numbers for Canada and Venezuela from [\[link\]](#) which describes how many barrels of oil and tons of lumber the workers can produce. Use these numbers to answer the rest of this question.

- a. Draw a production possibilities frontier for each country. Assume there are 100 workers in each country. Canadians and Venezuelans desire both oil and lumber. Canadians want at least 2,000 tons of lumber. Mark a point on their production possibilities where they can get at least 3,000 tons.
- b. Assume that the Canadians specialize completely because they figured out they have a comparative advantage in lumber. They are willing to give up 1,000 tons of lumber. How much oil should they ask for in return for this lumber to be as well off as they were

with no trade? How much should they ask for if they want to gain from trading with Venezuela? *Note:* We can think of this “ask” as the relative price or trade price of lumber.

- c. Is the Canadian “ask” you identified in (b) also beneficial for Venezuelans? Use the production possibilities frontier graph for Venezuela to show that Venezuelans can gain from trade.

Exercise:

Problem:

In [\[link\]](#), is there an “ask” where Venezuelans may say “no thank you” to trading with Canada?

References

Bernstein, William J. *A Splendid Exchange: How Trade Shaped the World*. Atlantic Monthly Press. New York. 2008.

Intra-industry Trade between Similar Economies

By the end of this section, you will be able to:

- Identify at least two advantages of intra-industry trading
- Explain the relationship between economies of scale and intra-industry trade

Absolute and comparative advantages explain a great deal about patterns of global trade. For example, they help to explain the patterns noted at the start of this chapter, like why you may be eating fresh fruit from Chile or Mexico, or why lower productivity regions like Africa and Latin America are able to sell a substantial proportion of their exports to higher productivity regions like the European Union and North America. Comparative advantage, however, at least at first glance, does not seem especially well-suited to explain other common patterns of international trade.

The Prevalence of Intra-industry Trade between Similar Economies

The theory of comparative advantage suggests that trade should happen between economies with large differences in opportunity costs of production. Roughly half of all world trade involves shipping goods between the fairly similar high-income economies of the United States, Canada, the European Union, Japan, Mexico, and China (see [\[link\]](#)).

Country	U.S. Exports Go to ...	U.S. Imports Come from ...
European Union	19.0%	21.0%

Country	U.S. Exports Go to ...	U.S. Imports Come from ...
Canada	22.0%	14.0%
Japan	4.0%	6.0%
Mexico	15.0%	13.0%
China	8.0%	20.0%

Where U.S. Exports Go and U.S. Imports Originate (2015)(Source: https://www.census.gov/foreign-trade/Press-Release/current_press_release/ft900.pdf)

Moreover, the theory of comparative advantage suggests that each economy should specialize to a degree in certain products, and then exchange those products. A high proportion of trade, however, is **intra-industry trade**—that is, trade of goods within the same industry from one country to another. For example, the United States produces and exports autos and imports autos. [\[link\]](#) shows some of the largest categories of U.S. exports and imports. In all of these categories, the United States is both a substantial exporter and a substantial importer of goods from the same industry. In 2014, according to the Bureau of Economic Analysis, the United States exported \$159 billion worth of autos, and imported \$327 billion worth of autos. About 60% of U.S. trade and 60% of European trade is intra-industry trade.

Some U.S. Exports	Quantity of Exports (\$ billions)	Quantity of Imports (\$ billions)
--------------------------	--	--

Some U.S. Exports	Quantity of Exports (\$ billions)	Quantity of Imports (\$ billions)
Autos	\$146	\$327
Food and beverages	\$144	\$126
Capital goods	\$550	\$551
Consumer goods	\$199	\$558
Industrial supplies	\$507	\$665
Other transportation	\$45	\$55

Some Intra-Industry U.S. Exports and Imports in 2014(Source: <http://www.bea.gov/newsreleases/international/trade/tradnewsrelease.htm>)

Why do similar high-income economies engage in intra-industry trade? What can be the economic benefit of having workers of fairly similar skills making cars, computers, machinery and other products which are then shipped across the oceans to and from the United States, the European Union, and Japan? There are two reasons: (1) The division of labor leads to learning, innovation, and unique skills; and (2) economies of scale.

Gains from Specialization and Learning

Consider the category of machinery, where the U.S. economy has considerable intra-industry trade. Machinery comes in many varieties, so the United States may be exporting machinery for manufacturing with wood, but importing machinery for photographic processing. The underlying reason why a country like the United States, Japan, or Germany

produces one kind of machinery rather than another is usually not related to U.S., German, or Japanese firms and workers having generally higher or lower skills. It is just that, in working on very specific and particular products, firms in certain countries develop unique and different skills.

Specialization in the world economy can be very finely split. In fact, recent years have seen a trend in international trade called **splitting up the value chain**. The **value chain** describes how a good is produced in stages. As indicated in the beginning of the chapter, the production of the iPhone involves the design and engineering of the phone in the United States, parts supplied from Korea, the assembly of the parts in China, and the advertising and marketing done in the United States. Thanks in large part to improvements in communication technology, sharing information, and transportation, it has become easier to split up the value chain. Instead of production in a single large factory, all of these steps can be split up among different firms operating in different places and even different countries. Because firms split up the value chain, international trade often does not involve whole finished products like automobiles or refrigerators being traded between nations. Instead, it involves shipping more specialized goods like, say, automobile dashboards or the shelving that fits inside refrigerators. Intra-industry trade between similar countries produces economic gains because it allows workers and firms to learn and innovate on particular products—and often to focus on very particular parts of the value chain.

Note:

Visit this [website](#) for some interesting information about the assembly of the iPhone.

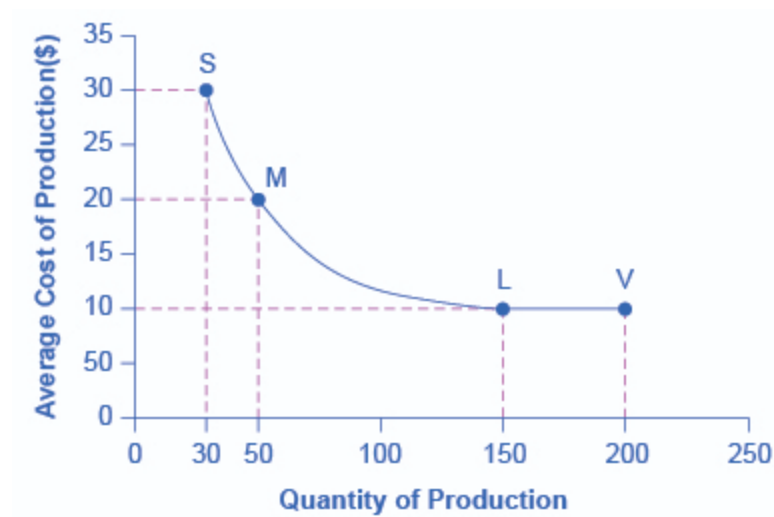


Economies of Scale, Competition, Variety

A second broad reason that intra-industry trade between similar nations produces economic gains involves economies of scale. The concept of economies of scale, as introduced in [Cost and Industry Structure](#), means that as the scale of output goes up, average costs of production decline—at least up to a point. [\[link\]](#) illustrates economies of scale for a plant producing toaster ovens. The horizontal axis of the figure shows the quantity of production by a certain firm or at a certain manufacturing plant. The vertical axis measures the average cost of production. Production plant S produces a small level of output at 30 units and has an average cost of production of \$30 per toaster oven. Plant M produces at a medium level of output at 50 units, and has an average cost of production of \$20 per toaster oven. Plant L produces 150 units of output with an average cost of production of only \$10 per toaster oven. Although plant V can produce 200 units of output, it still has the same unit cost as Plant L.

In this example, a small or medium plant, like S or M, will not be able to compete in the market with a large or a very large plant like L or V, because the firm that operates L or V will be able to produce and sell their output at a lower price. In this example, economies of scale operate up to point L, but beyond point L to V, the additional scale of production does not continue to reduce average costs of production.

Economies of Scale



Production Plant S, has an average cost of

production of \$30 per toaster oven.
Production plant M has an average cost of
production of \$20 per toaster oven.
Production plant L has an average cost of
production of only \$10 per toaster oven.
Production plant V would still have an
average cost of production of \$10 per
toaster oven. Thus, production plant M can
produce toaster ovens more cheaply than
plant S because of economies of scale, and
plants L or V can produce more cheaply
than S or M because of economies of scale.
However, the economies of scale end at an
output level of 150. Plant V, despite being
larger, cannot produce more cheaply on
average than plant L.

The concept of economies of scale becomes especially relevant to international trade when it enables one or two large producers to supply the entire country. For example, a single large automobile factory could probably supply all the cars purchased in a smaller economy like the United Kingdom or Belgium in a given year. However, if a country has only one or two large factories producing cars, and no international trade, then consumers in that country would have relatively little choice between kinds of cars (other than the color of the paint and other nonessential options). Little or no competition will exist between different car manufacturers.

International trade provides a way to combine the lower average production costs that come from economies of scale and still have competition and variety for consumers. Large automobile factories in different countries can make and sell their products around the world. If the U.S. automobile market was made up of only General Motors, Ford, and Chrysler, the level of competition and consumer choice would be quite a lot lower than when U.S. carmakers must face competition from Toyota, Honda, Suzuki, Fiat, Mitsubishi, Nissan, Volkswagen, Kia, Hyundai, BMW, Subaru, and others.

Greater competition brings with it innovation and responsiveness to what consumers want. America's car producers make far better cars now than they did several decades ago, and much of the reason is competitive pressure, especially from East Asian and European carmakers.

Dynamic Comparative Advantage

The sources of gains from intra-industry trade between similar economies—namely, the learning that comes from a high degree of specialization and splitting up the value chain and from economies of scale—do not contradict the earlier theory of comparative advantage. Instead, they help to broaden the concept.

In intra-industry trade, the level of worker productivity is not determined by climate or geography. It is not even determined by the general level of education or skill. Instead, the level of worker productivity is determined by how firms engage in specific learning about specialized products, including taking advantage of economies of scale. In this vision, comparative advantage can be dynamic—that is, it can evolve and change over time as new skills are developed and as the value chain is split up in new ways. This line of thinking also suggests that countries are not destined to have the same comparative advantage forever, but must instead be flexible in response to ongoing changes in comparative advantage.

Key Concepts and Summary

A large share of global trade happens between high-income economies that are quite similar in having well-educated workers and advanced technology. These countries practice intra-industry trade, in which they import and export the same products at the same time, like cars, machinery, and computers. In the case of intra-industry trade between economies with similar income levels, the gains from trade come from specialized learning in very particular tasks and from economies of scale. Splitting up the value chain means that several stages of producing a good take place in different countries around the world.

Self-Check Questions

Exercise:

Problem:

How can there be any economic gains for a country from both importing and exporting the same good, like cars?

Solution:

There are a number of possible advantages of intra-industry trade. Both nations can take advantage of extreme specialization and learning in certain kinds of cars with certain traits, like gas-efficient cars, luxury cars, sport-utility vehicles, higher- and lower-quality cars, and so on. Moreover, nations can take advantage of economies of scale, so that large companies will compete against each other across international borders, providing the benefits of competition and variety to customers. This same argument applies to trade between U.S. states, where people often buy products made by people of other states, even though a similar product is made within the boundaries of their own state. All states—and all countries—can benefit from this kind of competition and trade.

Exercise:

Problem:

[\[link\]](#) shows how the average costs of production for semiconductors (the “chips” in computer memories) change as the quantity of semiconductors built at that factory increases.

- a. Based on these data, sketch a curve with quantity produced on the horizontal axis and average cost of production on the vertical axis. How does the curve illustrate economies of scale?
- b. If the equilibrium quantity of semiconductors demanded is 90,000, can this economy take full advantage of economies of scale? What about if quantity demanded is 70,000?

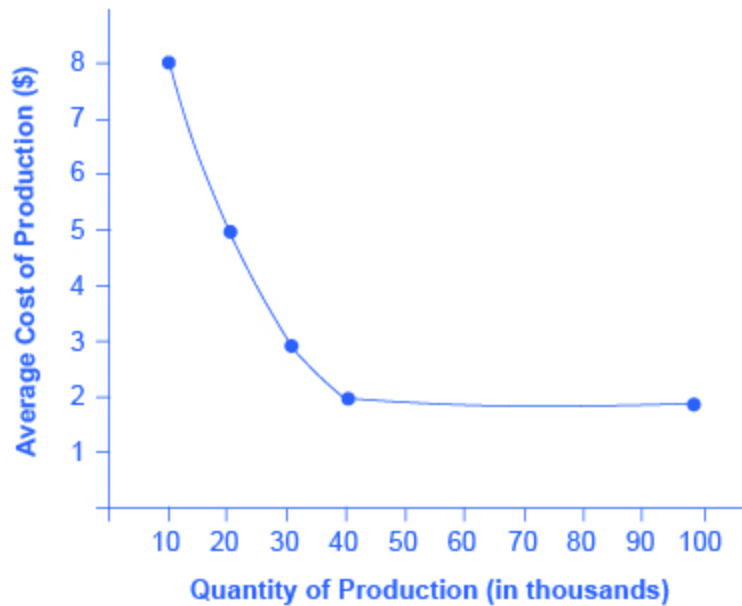
semiconductors? 50,000 semiconductors? 30,000 semiconductors?

- c. Explain how international trade could make it possible for even a small economy to take full advantage of economies of scale, while also benefiting from competition and the variety offered by several producers.

Quantity of Semiconductors	Average Total Cost
10,000	\$8 each
20,000	\$5 each
30,000	\$3 each
40,000	\$2 each
100,000	\$2 each

Solution:

(a) Start by plotting the points on a sketch diagram and then drawing a line through them. The following figure illustrates the average costs of production of semiconductors.



The curve illustrates economies of scale by showing that as the scale increases—that is, as production at this particular factory goes up—the average cost of production declines. The economies of scale exist up to an output of 40,000 semiconductors; at higher outputs, the average cost of production does not seem to decline any further.

(b) At any quantity demanded above 40,000, this economy can take full advantage of economies of scale; that is, it can produce at the lowest cost per unit. Indeed, if the quantity demanded was quite high, like 500,000, then there could be a number of different factories all taking full advantage of economies of scale and competing with each other. If the quantity demanded falls below 40,000, then the economy by itself, without foreign trade, cannot take full advantage of economies of scale.

(c) The simplest answer to this question is that the small country could have a large enough factory to take full advantage of economies of scale, but then export most of the output. For semiconductors, countries like Taiwan and Korea have recently fit this description. Moreover, this country could also import semiconductors from other countries which also have large factories, thus getting the benefits of

competition and variety. A slightly more complex answer is that the country can get these benefits of economies of scale without producing semiconductors, but simply by buying semiconductors made at low cost around the world. An economy, especially a smaller country, may well end up specializing and producing a few items on a large scale, but then trading those items for other items produced on a large scale, and thus gaining the benefits of economies of scale by trade, as well as by direct production.

Review Questions

Exercise:

Problem: What is intra-industry trade?

Exercise:

Problem:

What are the two main sources of economic gains from intra-industry trade?

Exercise:

Problem: What is splitting up the value chain?

Critical Thinking Questions

Exercise:

Problem:

Does intra-industry trade contradict the theory of comparative advantage?

Exercise:

Problem: Do consumers benefit from intra-industry trade?

Exercise:**Problem:**

Why might intra-industry trade seem surprising from the point of view of comparative advantage?

Problems**Exercise:****Problem:**

From earlier chapters you will recall that technological change shifts the average cost curves. Draw a graph showing how technological change could influence intra-industry trade.

Exercise:**Problem:**

Consider two countries: South Korea and Taiwan. Taiwan can produce one million mobile phones per day at the cost of \$10 per phone and South Korea can produce 50 million mobile phones at \$5 per phone. Assume these phones are the same type and quality and there is only one price. What is the minimum price at which both countries will engage in trade?

References

U.S. Census Bureau. 2015. "U.S. International Trade in Goods and Services: December 2014." Accessed April 13, 2015.
<http://www.bea.gov/newsreleases/international/trade/2015/pdf/trad1214.pdf>
.

U.S. Census Bureau. U.S. Bureau of Economic Analysis. 2015. "U.S. International Trade in Goods and Services February 2015." Accessed April

10, 2015. https://www.census.gov/foreign-trade/Press-Release/current_press_release/ft900.pdf.

Vernengo, Matias. “What Do Undergraduates Really Need to Know About Trade and Finance?” in *Political Economy and Contemporary Capitalism: Radical Perspectives on Economic Theory and Policy*, ed. Ron Baiman, Heather Boushey, and Dawn Saunders. M. E. Sharpe Inc, 2000. Armonk. 177-183.

Glossary

intra-industry trade

international trade of goods within the same industry

splitting up the value chain

many of the different stages of producing a good happen in different geographic locations

value chain

how a good is produced in stages

The Benefits of Reducing Barriers to International Trade

By the end of this section, you will be able to:

- Explain tariffs as barriers to trade
- Identify at least two benefits of reducing barriers to international trade

Tariffs are taxes that governments place on imported goods for a variety of reasons. Some of these reasons include protecting sensitive industries, for humanitarian reasons, and protecting against dumping. Traditionally, tariffs were used simply as a political tool to protect certain vested economic, social, and cultural interests. The World Trade Organization (WTO) is committed to lowering barriers to trade. The world's nations meet through the WTO to negotiate how they can reduce barriers to trade, such as tariffs. WTO negotiations happen in "rounds," where all countries negotiate one agreement to encourage trade, take a year or two off, and then start negotiating a new agreement. The current round of negotiations is called the Doha Round because it was officially launched in Doha, the capital city of Qatar, in November 2001. In 2009, economists from the World Bank summarized recent research and found that the Doha round of negotiations would increase the size of the world economy by \$160 billion to \$385 billion per year, depending on the precise deal that ended up being negotiated.

In the context of a global economy that currently produces more than \$30 trillion of goods and services each year, this amount is not huge: it is an increase of 1% or less. But before dismissing the gains from trade too quickly, it is worth remembering two points.

- First, a gain of a few hundred billion dollars is enough money to deserve attention! Moreover, remember that this increase is not a one-time event; it would persist each year into the future.
- Second, the estimate of gains may be on the low side because some of the gains from trade are not measured especially well in economic statistics. For example, it is difficult to measure the potential advantages to consumers of having a variety of products available and a greater degree of competition among producers. Perhaps the most important unmeasured factor is that trade between countries, especially

when firms are splitting up the value chain of production, often involves a transfer of knowledge that can involve skills in production, technology, management, finance, and law.

Low-income countries benefit more from trade than high-income countries do. In some ways, the giant U.S. economy has less need for international trade, because it can already take advantage of internal trade within its economy. However, many smaller national economies around the world, in regions like Latin America, Africa, the Middle East, and Asia, have much more limited possibilities for trade inside their countries or their immediate regions. Without international trade, they may have little ability to benefit from comparative advantage, slicing up the value chain, or economies of scale. Moreover, smaller economies often have fewer competitive firms making goods within their economy, and thus firms have less pressure from other firms to provide the goods and prices that consumers want.

The economic gains from expanding international trade are measured in hundreds of billions of dollars, and the gains from international trade as a whole probably reach well into the trillions of dollars. The potential for gains from trade may be especially high among the smaller and lower-income countries of the world.

Note:

Visit this [website](#) for a list of some benefits of trade.



From Interpersonal to International Trade

Most people find it easy to believe that they, personally, would not be better off if they tried to grow and process all of their own food, to make all of their own clothes, to build their own cars and houses from scratch, and so on. Instead, we all benefit from living in economies where people and firms can specialize and trade with each other.

The benefits of trade do not stop at national boundaries, either. Earlier we explained that the division of labor could increase output for three reasons: (1) workers with different characteristics can specialize in the types of production where they have a comparative advantage; (2) firms and workers who specialize in a certain product become more productive with learning and practice; and (3) economies of scale. These three reasons apply from the individual and community level right up to the international level. If it makes sense to you that interpersonal, intercommunity, and interstate trade offer economic gains, it should make sense that international trade offers gains, too.

International trade currently involves about \$20 trillion worth of goods and services moving around the globe. Any economic force of that size, even if it confers overall benefits, is certain to cause disruption and controversy. This chapter has only made the case that trade brings economic benefits. Other chapters discuss, in detail, the public policy arguments over whether to restrict international trade.

Note:

It's Apple's (Global) iPhone

Apple Corporation uses a global platform to produce the iPhone. Now that you understand the concept of comparative advantage, you can see why the engineering and design of the iPhone is done in the United States. The United States has built up a comparative advantage over the years in designing and marketing products, and sacrifices fewer resources to design high-tech devices relative to other countries. China has a comparative advantage in assembling the phone due to its large skilled labor force. Korea has a comparative advantage in producing components. Korea focuses its production by increasing its scale, learning better ways to produce screens and computer chips, and uses innovation to lower average

costs of production. Apple, in turn, benefits because it can purchase these quality products at lower prices. Put the global assembly line together and you have the device with which we are all so familiar.

Key Concepts and Summary

Tariffs are placed on imported goods as a way of protecting sensitive industries, for humanitarian reasons, and for protection against dumping. Traditionally, tariffs were used as a political tool to protect certain vested economic, social, and cultural interests. The WTO has been, and continues to be, a way for nations to meet and negotiate through barriers to trade. The gains of international trade are very large, especially for smaller countries, but are beneficial to all.

Self-Check Question

Exercise:

Problem:

If the removal of trade barriers is so beneficial to international economic growth, why would a nation continue to restrict trade on some imported or exported products?

Solution:

A nation might restrict trade on imported products to protect an industry that is important for national security. For example, nation X and nation Y may be geopolitical rivals, each with ambitions of increased political and economic strength. Even if nation Y has comparative advantage in the production of missile defense systems, it is unlikely that nation Y would seek to export those goods to nation X. It is also the case that, for some nations, the production of a particular good is a key component of national identity. In Japan, the production of rice is culturally very important. It may be difficult for Japan to

import rice from a nation like Vietnam, even if Vietnam has a comparative advantage in rice production.

Review Question

Exercise:

Problem:

Are the gains from international trade more likely to be relatively more important to large or small countries?

Critical Thinking Questions

Exercise:

Problem:

In World Trade Organization meetings, what do you think low-income countries lobby for?

Exercise:

Problem:

Why might a low-income country put up barriers to trade, such as tariffs on imports?

Exercise:

Problem:

Can a nation's comparative advantage change over time? What factors would make it change?

Problems

Exercise:

Problem:

If trade increases world GDP by 1% per year, what is the global impact of this increase over 10 years? How does this increase compare to the annual GDP of a country like Sri Lanka? Discuss. *Hint:* To answer this question, here are steps you may want to consider. Go to the World Development Indicators (online) published by the World Bank. Find the current level of World GDP in constant international dollars. Also, find the GDP of Sri Lanka in constant international dollars. Once you have these two numbers, compute the amount the additional increase in global incomes due to trade and compare that number to Sri Lanka's GDP.

References

World Trade Organization. "The Doha Round." Accessed October 2013. http://www.wto.org/english/tratop_e/dda_e/dda_e.htm.

The World Bank. "Data: World Development Indicators." Accessed October 2013. <http://data.worldbank.org/data-catalog/world-development-indicators>.

Glossary

tariffs

taxes that governments place on imported goods

Introduction to Globalization and Protectionism

class="introduction"

Flat Screen Competition

The market for flat-panel displays in the United States is huge. The manufacturers of flat screens in the United States must compete against manufacturers from around the world. (Credit: modification of work by “Jemimus”/Flickr r Creative Commons)

**Note:****What's the Downside of Protection?**

Governments are motivated to limit and alter market outcomes for political or social ends. While governments can limit the rise in prices of some products, they cannot control how much people want to buy or how much firms are willing to sell. The laws of demand and supply still hold. Trade policy is an example where regulations can redirect economic forces, but it cannot stop them from manifesting themselves elsewhere.

Flat-panel displays, the displays for laptop computers, tablets, and flat screen televisions, are an example of such an enduring principle. In the early 1990s, the vast majority of flat-panel displays used in U.S.-manufactured laptops were imported, primarily from Japan. The small but politically powerful U.S. flat-panel-display industry filed a dumping complaint with the Commerce Department. They argued that Japanese firms were selling displays at “less than fair value,” which made it difficult for U.S. firms to compete. This argument for trade protection is referred to as anti-dumping. Other arguments for protection in this complaint included national security. After a preliminary determination by the Commerce Department that the Japanese firms were dumping, the U.S. International

Trade Commission imposed a 63% dumping margin (or tax) on the import of flat-panel displays. Was this a successful exercise of U.S. trade policy? See what you think after reading the chapter.

Note:

Introduction to Globalism and Protectionism

In this chapter, you will learn about:

- Protectionism: An Indirect Subsidy from Consumers to Producers
- International Trade and Its Effects on Jobs, Wages, and Working Conditions
- Arguments in Support of Restricting Imports
- How Trade Policy Is Enacted: Globally, Regionally, and Nationally
- The Tradeoffs of Trade Policy

The world has become more connected on multiple levels, especially economically. In 1970, imports and exports made up 11% of U.S. GDP, while now they make up 32%. However, the United States, due to its size, is less internationally connected than most countries. For example, according to the World Bank, 97% of Botswana's economic activity is connected to trade. This chapter explores trade policy—the laws and strategies a country uses to regulate international trade. This topic is not without controversy.

As the world has become more globally connected, firms and workers in high-income countries like the United States, Japan, or the nations of the European Union, perceive a competitive threat from firms in medium-income countries like Mexico, China, or South Africa, that have lower costs of living and therefore pay lower wages. Firms and workers in low-income countries fear that they will suffer if they must compete against more productive workers and advanced technology in high-income countries.

On a different tack, some environmentalists worry that multinational firms may evade environmental protection laws by moving their production to

countries with loose or nonexistent pollution standards, trading a clean environment for jobs. Some politicians worry that their country may become overly dependent on key imported products, like oil, which in a time of war could threaten national security. All of these fears influence governments to reach the same basic policy conclusion: to protect national interests, whether businesses, jobs, or security, imports of foreign products should be restricted. This chapter analyzes such arguments. First, however, it is essential to learn a few key concepts and understand how the demand and supply model applies to international trade.

Protectionism: An Indirect Subsidy from Consumers to Producers

By the end of this section, you will be able to:

- Explain protectionism and its three main forms
- Analyze protectionism through concepts of demand and supply, noting its effects on equilibrium
- Calculate the effects of trade barriers

When a government legislates policies to reduce or block international trade it is engaging in **protectionism**. Protectionist policies often seek to shield domestic producers and domestic workers from foreign competition. Protectionism takes three main forms: tariffs, import quotas, and nontariff barriers.

Recall from [International Trade](#) that tariffs are taxes imposed on imported goods and services. They make imports more expensive for consumers, discouraging imports. For example, in recent years large, flat-screen televisions imported from China have faced a 5% tariff rate.

Another way to control trade is through **import quotas**, which are numerical limitations on the quantity of products that can be imported. For instance, during the early 1980s, the Reagan Administration imposed a quota on the import of Japanese automobiles. In the 1970s, many developed countries, including the United States, found themselves with declining textile industries. Textile production does not require highly skilled workers, so producers were able to set up lower-cost factories in developing countries. In order to “manage” this loss of jobs and income, the developed countries established an international Multifiber Agreement that essentially divided up the market for textile exports between importers and the remaining domestic producers. The agreement, which ran from 1974 to 2004, specified the exact quota of textile imports that each developed country would accept from each low-income country. A similar story exists for sugar imports into the United States, which are still governed by quotas.

Nontariff barriers are all the other ways that a nation can draw up rules, regulations, inspections, and paperwork to make it more costly or difficult to import products. A rule requiring certain safety standards can limit

imports just as effectively as high tariffs or low import quotas, for instance. There are also nontariff barriers in the form of “rules-of-origin” regulations—these rules describe the “Made in Country X” label as the one in which the last substantial change in the product took place. A manufacturer wishing to evade import restrictions may try to change the production process so that the last big change in the product happens in his or her own country. For example, certain textiles are made in the United States, shipped to other countries, combined with textiles made in those other countries to make apparel—and then re-exported back to the United States for a final assembly, to escape paying tariffs or to obtain a “Made in the USA” label.

Despite import quotas, tariffs, and nontariff barriers, the share of apparel sold in the United States that is imported rose from about half in 1999 to about three-quarters today. The U.S. Bureau of Labor Statistics (BLS), estimated the number of U.S. jobs in textiles and apparel fell from 666,360 in 2007 to 385,240 in 2012, a 42% decline. Even more U.S. textile industry jobs would have been lost without tariffs, however, domestic jobs that are saved by import quotas come at a cost. Because textile and apparel protectionism adds to the costs of imports, consumers end up paying billions of dollars more for clothing each year.

When the United States eliminates trade barriers in one area, consumers spend the money they save on that product elsewhere in the economy—so there is no overall loss of jobs for the economy as a whole. Of course, workers in some of the poorest countries of the world who would otherwise have jobs producing textiles, would gain considerably if the United States reduced its barriers to trade in textiles. That said, there are good reasons to be wary about reducing barriers to trade. The 2012 and 2013 Bangladeshi fires in textile factories, which resulted in a horrific loss of life, present complications that our simplified analysis in the chapter will not capture.

Realizing the compromises between nations that come about due to trade policy, many countries came together in 1947 to form the General Agreement on Tariffs and Trade (GATT). (We’ll cover the GATT in more detail later in the chapter.) This agreement has since been superseded by the **World Trade Organization (WTO)**, whose membership includes about 150 nations and most of the economies of the world. It is the primary

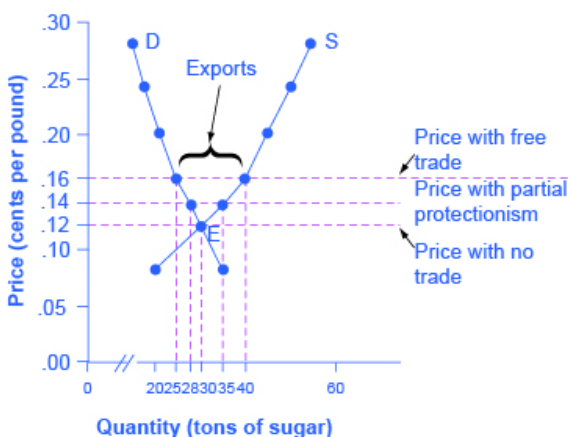
international mechanism through which nations negotiate their trade rules—including rules about tariffs, quotas, and nontariff barriers. The next section examines the results of such protectionism and develops a simple model to show the impact of trade policy.

Demand and Supply Analysis of Protectionism

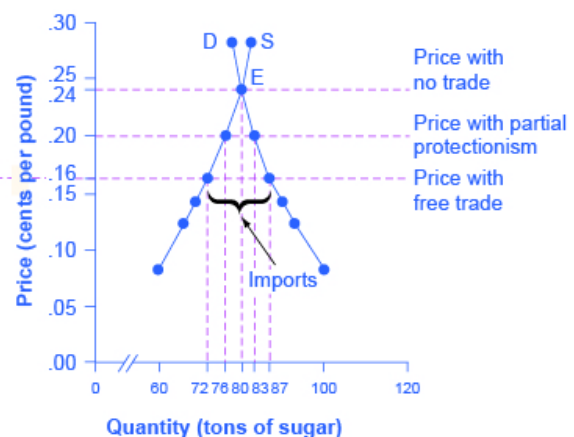
To the non-economist, restricting imports may appear to be nothing more than taking sales from foreign producers and giving them to domestic producers. Other factors are at work, however, because firms do not operate in a vacuum. Instead, firms sell their products either to consumers or to other firms (if they are business suppliers), who are also affected by the trade barriers. A demand and supply analysis of protectionism shows that it is not just a matter of domestic gains and foreign losses, but a policy that imposes substantial domestic costs as well.

Consider two countries, Brazil and the United States, who produce sugar. Each country has a domestic supply and demand for sugar, as detailed in [\[link\]](#) and illustrated in [\[link\]](#). In Brazil, without trade, the equilibrium price of sugar is 12 cents per pound and the equilibrium output is 30 tons. When there is no trade in the United States, the equilibrium price of sugar is 24 cents per pound and the equilibrium quantity is 80 tons. These equilibrium points are labeled with the point E.

The Sugar Trade between Brazil and the United States



(a) Brazil



(b) The United States

Before trade, the equilibrium price of sugar in Brazil is 12 cents a

pound and for 24 cents per pound in the United States. When trade is allowed, businesses will buy cheap sugar in Brazil and sell it in the United States. This will result in higher prices in Brazil and lower prices in the United States. Ignoring transaction costs, prices should converge to 16 cents per pound, with Brazil exporting 15 tons of sugar and the United States importing 15 tons of sugar. If trade is only partly open between the countries, it will lead to an outcome between the free-trade and no-trade possibilities.

Price	Brazil: Quantity Supplied (tons)	Brazil: Quantity Demanded (tons)	U.S.: Quantity Supplied (tons)	U.S.: Quantity Demanded (tons)
8 cents	20	35	60	100
12 cents	30	30	66	93
14 cents	35	28	69	90
16 cents	40	25	72	87
20 cents	45	21	76	83

Price	Brazil: Quantity Supplied (tons)	Brazil: Quantity Demanded (tons)	U.S.: Quantity Supplied (tons)	U.S.: Quantity Demanded (tons)
24 cents	50	18	80	80
28 cents	55	15	82	78

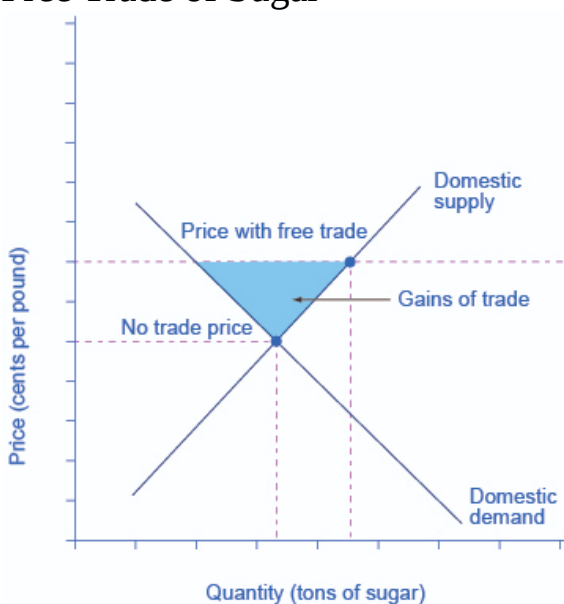
The Sugar Trade between Brazil and the United States

If international trade between Brazil and the United States now becomes possible, profit-seeking firms will spot an opportunity: buy sugar cheaply in Brazil, and sell it at a higher price in the United States. As sugar is shipped from Brazil to the United States, the quantity of sugar produced in Brazil will be greater than Brazilian consumption (with the extra production being exported), and the amount produced in the United States will be less than the amount of U.S. consumption (with the extra consumption being imported). Exports to the United States will reduce the supply of sugar in Brazil, raising its price. Imports into the United States will increase the supply of sugar, lowering its price. When the price of sugar is the same in both countries, there is no incentive to trade further. As [\[link\]](#) shows, the equilibrium with trade occurs at a price of 16 cents per pound. At that price, the sugar farmers of Brazil supply a quantity of 40 tons, while the consumers of Brazil buy only 25 tons.

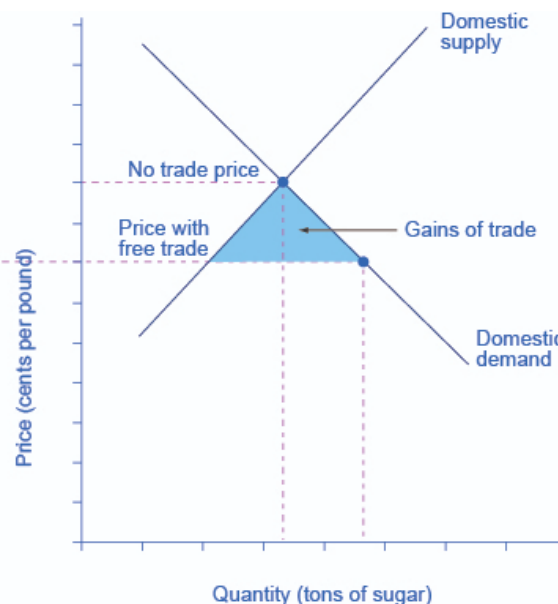
The extra 15 tons of sugar production, shown by the horizontal gap between the demand curve and the supply curve in Brazil, is exported to the United States. In the United States, at a price of 16 cents, the farmers produce a quantity of 72 tons and consumers demand a quantity of 87 tons. The excess demand of 15 tons by American consumers, shown by the horizontal gap between demand and domestic supply at the price of 16 cents, is supplied by imported sugar. Free trade typically results in income distribution effects, but the key is to recognize the overall gains from trade, as shown in [\[link\]](#). Building on the concepts outlined in [Demand and](#)

[Supply](#) and [Demand](#), [Supply, and Efficiency](#) in terms of consumer and producer surplus, [\[link\]](#) (a) shows that producers in Brazil gain by selling more sugar at a higher price, while [\[link\]](#) (b) shows consumers in the United States benefit from the lower price and greater availability of sugar. Consumers in Brazil are worse off (compare their no-trade consumer surplus with the free-trade consumer surplus) and U.S. producers of sugar are worse off. There are gains from trade—an increase in social surplus in each country. That is, both the United States and Brazil are better off than they would be without trade. The following Clear It Up feature explains how trade policy can influence low-income countries.

Free Trade of Sugar



(a) Brazil



(b) The United States

Free trade results in gains from trade. Total surplus increases in both countries. However, there are clear income distribution effects.

Note:

Visit this [website](#) to read more about the global sugar trade.

**Note:**

Why are there low-income countries?

Why are the poor countries of the world poor? There are a number of reasons, but one of them will surprise you: the trade policies of the high-income countries. Following is a stark review of social priorities which has been widely publicized by the international aid organization, Oxfam International.

High-income countries of the world—primarily the United States, Canada, countries of the European Union, and Japan—subsidize their domestic farmers collectively by about \$360 billion per year. By contrast, the total amount of foreign aid from these same high-income countries to the poor countries of the world is about \$70 billion per year, or less than 20% of the farm subsidies. Why does this matter?

It matters because the support of farmers in high-income countries is devastating to the livelihoods of farmers in low-income countries. Even when their climate and land are well-suited to products like cotton, rice, sugar, or milk, farmers in low-income countries find it difficult to compete. Farm subsidies in the high-income countries cause farmers in those countries to increase the amount they produce. This increase in supply drives down world prices of farm products below the costs of production. As Michael Gerson of the *Washington Post* describes it: “[T]he effects in the cotton-growing regions of West Africa are dramatic . . . keep[ing] millions of Africans on the edge of malnutrition. In some of the poorest countries on Earth, cotton farmers are some of the poorest people, earning about a dollar a day. . . . Who benefits from the current system of subsidies? About 20,000 American cotton producers, with an average annual income of more than \$125,000.”

As if subsidies were not enough, often, the high-income countries block agricultural exports from low-income countries. In some cases, the situation gets even worse when the governments of high-income countries, having bought and paid for an excess supply of farm products, give away those products in poor countries and drive local farmers out of business altogether.

For example, shipments of excess milk from the European Union to Jamaica have caused great hardship for Jamaican dairy farmers. Shipments of excess rice from the United States to Haiti drove thousands of low-income rice farmers in Haiti out of business. The opportunity costs of protectionism are not paid just by domestic consumers, but also by foreign producers—and for many agricultural products, those foreign producers are the world's poor.

Now, let's look at what happens with protectionism. U.S. sugar farmers are likely to argue that, if only they could be protected from sugar imported from Brazil, the United States would have higher domestic sugar production, more jobs in the sugar industry, and American sugar farmers would receive a higher price. If the United States government sets a high-enough tariff on imported sugar, or sets an import quota at zero, the result will be that the quantity of sugar traded between countries could be reduced to zero, and the prices in each country will return to the levels before trade was allowed.

Blocking only some trade is also possible. Suppose that the United States passed a sugar import quota of seven tons. The United States will import no more than seven tons of sugar, which means that Brazil can export no more than seven tons of sugar to the United States. As a result, the price of sugar in the United States will be 20 cents, which is the price where the quantity demanded is seven tons greater than the domestic quantity supplied. Conversely, if Brazil can export only seven tons of sugar, then the price of sugar in Brazil will be 14 cents per pound, which is the price where the domestic quantity supplied in Brazil is seven tons greater than domestic demand.

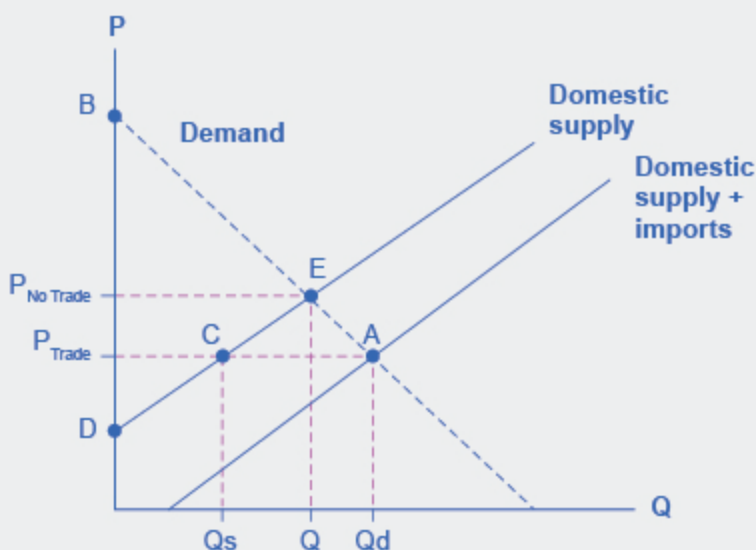
In general, when a country sets a low or medium tariff or import quota, the equilibrium price and quantity will be somewhere between no trade and completely free trade. The following Work It Out explores the impact of these trade barriers.

Note:

Effects of Trade Barriers

Let's look carefully at the effects of tariffs or quotas. If the U.S. government imposes a tariff or quota sufficient to eliminate trade with Brazil, two things occur: U.S. consumers pay a higher price and therefore buy a smaller quantity of sugar. U.S. producers obtain a higher price so they sell a larger quantity of sugar. The effects of a tariff on producers and consumers in the United States can be measured using two concepts developed in [Demand, Supply, and Efficiency](#): consumer surplus and producer surplus.

U.S. Sugar Supply and Demand



When there is free trade, the equilibrium is at point A. When there is no trade, the equilibrium is at point E.

Step 1. Look at [\[link\]](#), which shows a hypothetical version of the demand and supply of sugar in the United States.

Step 2. Note that the sugar market is in equilibrium at point A where Domestic Quantity Demanded (Q_d) = Quantity Supplied (Domestic Q_s + Imports from Brazil) at a price of P_{Trade} when there is free trade.

Step 3. Note, also, that imports are equal to the distance between points C and A.

Step 4. Recall that consumer surplus is the value a consumer gets beyond what they paid for when they buy a product. Graphically, it is the area under a demand curve but above the price. In this case, the consumer surplus in the United States is the area of the triangle formed by the points P_{Trade} , A, and B.

Step 5. Recall, also, that producer surplus is another name for profit—it is the income producers get above the cost of production, which is shown by the supply curve here. In this case, the producer surplus with trade is the area of the triangle formed by the points P_{trade} , C, and D.

Step 6. Suppose that the barriers to trade are imposed, imports are excluded, and the price rises to P_{NoTrade} . Look what happens to producer surplus and consumer surplus. At the higher price, the domestic quantity supplied increases from Q_s to Q at point E. Because producers are selling more quantity at a higher price, the producer surplus increases to the area of the triangle P_{NoTrade} , E, and D.

Step 7. Compare the areas of the two triangles and you will see the increase in the producer surplus.

Step 8. Examine the consumer surplus. Consumers are now paying a higher price to get a lower quantity (Q instead of Q_d). Their consumer surplus shrinks to the area of the triangle P_{NoTrade} , E, and B.

Step 9. Determine the net effect. The producer surplus increases by the area P_{trade} , C, E, P_{NoTrade} . The loss of consumer surplus, however, is larger. It is the area P_{trade} , A, E, P_{NoTrade} . In other words, consumers lose more than producers gain as a result of the trade barriers and the United States has a lower social surplus.

Who Benefits and Who Pays?

Using the demand and supply model, consider the impact of protectionism on producers and consumers in each of the two countries. For protected producers like U.S. sugar farmers, restricting imports is clearly positive. Without a need to face imported products, these producers are able to sell more, at a higher price. For consumers in the country with the protected good, in this case U.S. sugar consumers, restricting imports is clearly negative. They end up buying a lower quantity of the good and paying a higher price for what they do buy, compared to the equilibrium price and quantity without trade. The following Clear It Up feature considers why a country might outsource jobs even for a domestic product.

Note:

Why are Life Savers, an American product, not made in America?

Life Savers, the hard candy with the hole in the middle, were invented in 1912 by Clarence Crane in Cleveland, Ohio. Starting in the late 1960s and for 35 years afterward, 46 billion Life Savers a year, in 200 million rolls, were produced by a plant in Holland, Michigan. But in 2002, the Kraft Company announced that the Michigan plant would be closed and Life Saver production moved across the border to Montreal, Canada.

One reason is that Canadian workers are paid slightly less, especially in healthcare and insurance costs that are not linked to employment there.

Another main reason is that the United States government keeps the price of sugar high for the benefit of sugar farmers, with a combination of a government price floor program and strict quotas on imported sugar.

According to the Coalition for Sugar Reform, from 2009 to 2012, the price of refined sugar in the United States ranged from 64% to 92% higher than the world price. Life Saver production uses over 100 tons of sugar each day, because the candies are 95% sugar.

A number of other candy companies have also reduced U.S. production and expanded foreign production. Indeed, from 1997 to 2011, some 127,000 jobs in the sugar-using industries, or more than seven times the total employment in sugar production, were eliminated. While the candy industry is especially affected by the cost of sugar, the costs are spread more broadly. U.S. consumers pay roughly \$1 billion per year in higher food prices because of elevated sugar costs. Meanwhile, sugar producers in

low-income countries are driven out of business. Because of the sugar subsidies to domestic producers and the quotas on imports, they cannot sell their output profitably, or at all, in the United States market.

The fact that protectionism pushes up prices for consumers in the country enacting such protectionism is not always acknowledged openly, but it is not disputed. After all, if protectionism did not benefit domestic producers, there would not be much point in enacting such policies in the first place. Protectionism is simply a method of requiring consumers to subsidize producers. The subsidy is indirect, since it is paid by consumers through higher prices, rather than a direct subsidy paid by the government with money collected from taxpayers. But protectionism works like a subsidy, nonetheless. The American satirist Ambrose Bierce defined “tariff” this way in his 1911 book, *The Devil’s Dictionary*: “Tariff, n. A scale of taxes on imports, designed to protect the domestic producer against the greed of his consumer.”

The effect of protectionism on producers and consumers in the foreign country is complex. When an import quota is used to impose partial protectionism, the sugar producers of Brazil receive a lower price for the sugar they sell in Brazil—but a higher price for the sugar they are allowed to export to the United States. Indeed, notice that some of the burden of protectionism, paid by domestic consumers, ends up in the hands of foreign producers in this case. Brazilian sugar consumers seem to benefit from U.S. protectionism, because it reduces the price of sugar that they pay. On the other hand, at least some of these Brazilian sugar consumers also work as sugar farmers, so their incomes and jobs are reduced by protectionism. Moreover, if trade between the countries vanishes, Brazilian consumers would miss out on better prices for imported goods—which do not appear in our single-market example of sugar protectionism.

The effects of protectionism on foreign countries notwithstanding, protectionism requires domestic consumers of a product (consumers may include either households or other firms) to pay higher prices to benefit domestic producers of that product. In addition, when a country enacts protectionism, it loses the economic gains it would have been able to

achieve through a combination of comparative advantage, specialized learning, and economies of scale, concepts discussed in [International Trade](#).

Key Concepts and Summary

There are three tools for restricting the flow of trade: tariffs, import quotas, and nontariff barriers. When a country places limitations on imports from abroad, regardless of whether it uses tariffs, quotas, or nontariff barriers, it is said to be practicing protectionism. Protectionism will raise the price of the protected good in the domestic market, which causes domestic consumers to pay more, but domestic producers to earn more.

Self-Check Questions

Exercise:

Problem:

Explain how a tariff reduction causes an increase in the equilibrium quantity of imports and a decrease in the equilibrium price. *Hint:* Consider the [Work It Out](#) "Effects of Trade Barriers."

Solution:

This is the opposite case of the Work It Out feature. A reduced tariff is like a decrease in the cost of production, which is shown by a downward (or rightward) shift in the supply curve.

Exercise:

Problem:

Explain how a subsidy on agricultural goods like sugar adversely affects the income of foreign producers of imported sugar.

Solution:

A subsidy is like a reduction in cost. This shifts the supply curve down (or to the right), driving the price of sugar down. If the subsidy is large

enough, the price of sugar can fall below the cost of production faced by foreign producers, which means they will lose money on any sugar they produce and sell.

Review Questions

Exercise:

Problem:

Who does protectionism protect? What does it protect them from?

Exercise:

Problem:

Name and define three policy tools for enacting protectionism.

Exercise:

Problem:

How does protectionism affect the price of the protected good in the domestic market?

Critical Thinking Questions

Exercise:

Problem:

Show graphically that for any tariff, there is an equivalent quota that would give the same result. What would be the difference, then, between the two types of trade barriers? *Hint:* It is not something you can see from the graph.

Exercise:

Problem:

From the [Work It Out](#) "Effects of Trade Barriers," you can see that a tariff raises the price of imports. What is interesting is that the price rises by less than the amount of the tariff. Who pays the rest of the tariff amount? Can you show this graphically?

Problems**Exercise:****Problem:**

Assume two countries, Thailand (T) and Japan (J), have one good: cameras. The demand (d) and supply (s) for cameras in Thailand and Japan is described by the following functions:

Equation:

$$Q_d^T = 60 - P$$

Equation:

$$Q_s^T = -5 + \frac{1}{4}P$$

Equation:

$$Q_d^J = 80 - P$$

Equation:

$$Q_s^J = -10 + \frac{1}{2}P$$

P is the price measured in a common currency used in both countries, such as the Thai Baht.

- a. Compute the equilibrium price (P) and quantities (Q) in each country without trade.
- b. Now assume that free trade occurs. The free-trade price goes to 56.36 Baht. Who exports and imports cameras and in what quantities?

References

Bureau of Labor Statistics. "Industries at a Glance." Accessed December 31, 2013. <http://www.bls.gov/iag/>.

Oxfam International. Accessed January 6, 2014. <http://www.oxfam.org/>.

Glossary

import quotas

numerical limits on the quantity of products that can be imported

nontariff barriers

ways a nation can draw up rules, regulations, inspections, and paperwork to make it more costly or difficult to import products

protectionism

government policies to reduce or block imports

World Trade Organization (WTO)

organization that seeks to negotiate reductions in barriers to trade and to adjudicate complaints about violations of international trade policy; successor to the General Agreement on Tariffs and Trade (GATT)

International Trade and Its Effects on Jobs, Wages, and Working Conditions

By the end of this section, you will be able to:

- Discuss how international trade influences the job market
- Analyze the opportunity cost of protectionism
- Explain how international trade impacts wages, labor standards, and working conditions

In theory at least, imports might injure workers in several different ways: fewer jobs, lower wages, or poor working conditions. Let's consider these in turn.

Fewer Jobs?

In the early 1990s, the United States was negotiating the North American Free Trade Agreement (NAFTA) with Mexico, an agreement that reduced tariffs, import quotas, and nontariff barriers to trade between the United States, Mexico, and Canada. H. Ross Perot, a 1992 candidate for U.S. president, claimed, in prominent campaign arguments, that if the United States expanded trade with Mexico, there would be a “giant sucking sound” as U.S. employers relocated to Mexico to take advantage of lower wages. After all, average wages in Mexico were, at that time, about one-eighth of those in the United States. NAFTA passed Congress, President Bill Clinton signed it into law, and it took effect in 1995. For the next six years, the United States economy had some of the most rapid job growth and low unemployment in its history. Those who feared that open trade with Mexico would lead to a dramatic decrease in jobs were proven wrong.

This result was no surprise to economists. After all, the trend toward globalization has been going on for decades, not just since NAFTA. If trade did reduce the number of available jobs, then the United States should have been seeing a steady loss of jobs for decades. While the United States economy does experience rises and falls in unemployment rates—according to the Bureau of Labor Statistics, from spring 2008 to late 2009, the unemployment rate rose from 4.4% to 10%; it has since fallen back to 5.5% in spring 2015—the number of jobs is not falling over extended periods of

time. The number of U.S. jobs rose from 71 million in 1970 to 138 million in 2012.

Protectionism certainly saves jobs in the specific industry being protected but, for two reasons, it costs jobs in other unprotected industries. First, if consumers are paying higher prices to the protected industry, they inevitably have less money to spend on goods from other industries, and so jobs are lost in those other industries. Second, if the protected product is sold to other firms, so that other firms must now pay a higher price for a key input, then those firms will lose sales to foreign producers who do not need to pay the higher price. Lost sales translate into lost jobs. The hidden opportunity cost of using protectionism to save jobs in one industry is jobs sacrificed in other industries. This is why the United States International Trade Commission, in its study of barriers to trade, predicts that reducing trade barriers would not lead to an overall loss of jobs. Protectionism reshuffles jobs from industries without import protections to industries that are protected from imports, but it does not create more jobs.

Moreover, the costs of saving jobs through protectionism can be very high. A number of different studies have attempted to estimate the cost to consumers in higher prices per job saved through protectionism. [\[link\]](#) shows a sample of results, compiled by economists at the Federal Reserve Bank of Dallas. Saving a job through protectionism typically costs much more than the actual worker's salary. For example, a study published in 2002 compiled evidence that using protectionism to save an average job in the textile and apparel industry would cost \$199,000 per job saved. In other words, those workers could have been paid \$100,000 per year to be unemployed and the cost would only be half of what it is to keep them working in the textile and apparel industry. This result is not unique to textiles and apparel.

Industry Protected with Import Tariffs or Quotas	Annual Cost per Job Saved
Sugar	\$826,000
Polyethylene resins	\$812,000
Dairy products	\$685,000
Frozen concentrated orange juice	\$635,000
Ball bearings	\$603,000
Machine tools	\$479,000
Women's handbags	\$263,000
Glassware	\$247,000
Apparel and textiles	\$199,000
Rubber footwear	\$168,000
Women's nonathletic footwear	\$139,000

Cost to U.S. Consumers of Saving a Job through Protectionism(Source: Federal Reserve Bank of Dallas)

Why does it cost so much to save jobs through protectionism? The basic reason is that not all of the extra money paid by consumers because of tariffs or quotas goes to save jobs. For example, if tariffs are imposed on steel imports so that buyers of steel pay a higher price, U.S. steel companies earn greater profits, buy more equipment, pay bigger bonuses to managers, give pay raises to existing employees—and also avoid firing some additional workers. Only part of the higher price of protected steel goes toward saving jobs. Also, when an industry is protected, the economy as a whole loses the benefits of playing to its comparative advantage—in other

words, producing what it is best at. So, part of the higher price that consumers pay for protected goods is lost economic efficiency, which can be measured as another deadweight loss, like that discussed in [Labor and Financial Markets](#).

There's a bumper sticker that speaks to the threat some U.S. workers feel from imported products: "Buy American—Save U.S. Jobs." If the car were being driven by an economist, the sticker might declare: "Block Imports—Save Jobs for Some Americans, Lose Jobs for Other Americans, and Also Pay High Prices."

Trade and Wages

Even if trade does not reduce the number of jobs, it could affect wages. Here, it is important to separate issues about the average level of wages from issues about whether the wages of certain workers may be helped or hurt by trade.

Because trade raises the amount that an economy can produce by letting firms and workers play to their comparative advantage, trade will also cause the average level of wages in an economy to rise. Workers who can produce more will be more desirable to employers, which will shift the demand for their labor out to the right, and increase wages in the labor market. By contrast, barriers to trade will reduce the average level of wages in an economy.

However, even if trade increases the overall wage level, it will still benefit some workers and hurt others. Workers in industries that are confronted by competition from imported products may find that demand for their labor decreases and shifts back to the left, so that their wages decline with a rise in international trade. Conversely, workers in industries that benefit from selling in global markets may find that demand for their labor shifts out to the right, so that trade raises their wages.

Note:

View this [website](#) to read an article on the issues surrounding fair trade coffee.



One concern is that while globalization may be benefiting high-skilled, high-wage workers in the United States, it may also impose costs on low-skilled, low-wage workers. After all, high-skilled U.S. workers presumably benefit from increased sales of sophisticated products like computers, machinery, and pharmaceuticals in which the United States has a comparative advantage. Meanwhile, low-skilled U.S. workers must now compete against extremely low-wage workers worldwide for making simpler products like toys and clothing. As a result, the wages of low-skilled U.S. workers are likely to fall. There are, however, a number of reasons to believe that while globalization has helped some U.S. industries and hurt others, it has not focused its negative impact on the wages of low-skilled Americans. First, about half of U.S. trade is intra-industry trade. That means the U.S. trades similar goods with other high-wage economies like Canada, Japan, Germany, and the United Kingdom. For instance, in 2014 the U.S. exported over 2 million cars, from all the major automakers, and also imported several million cars from other countries.

Most U.S. workers in these industries have above-average skills and wages—and many of them do quite well in the world of globalization. Some evidence suggested that intra-industry trade between similar countries had a small impact on domestic workers but later evidence indicates that it all depends on how flexible the labor market is. In other words, the key is how flexible workers are in finding jobs in different industries. Trade on low-wage workers depends a lot on the structure of labor markets and indirect effects felt in other parts of the economy. For example, in the United States

and the United Kingdom, because labor market frictions are low, the impact of trade on low income workers is small.

Second, many low-skilled U.S. workers hold service jobs that cannot be replaced by imports from low-wage countries. For example, lawn care services or moving and hauling services or hotel maids cannot be imported from countries long distances away like China or Bangladesh. Competition from imported products is not the primary determinant of their wages.

Finally, while the focus of the discussion here is on wages, it is worth pointing out that low-wage U.S. workers suffer due to protectionism in all the industries—even those that they do not work in the U.S. For example, food and clothing are protected industries. These low-wage workers therefore pay higher prices for these basic necessities and as such their dollar stretches over fewer goods.

The benefits and costs of increased trade in terms of its effect on wages are not distributed evenly across the economy. However, the growth of international trade has helped to raise the productivity of U.S. workers as a whole—and thus helped to raise the average level of wages.

Labor Standards and Working Conditions

Workers in many low-income countries around the world labor under conditions that would be illegal for a worker in the United States. Workers in countries like China, Thailand, Brazil, South Africa, and Poland are often paid less than the United States minimum wage. For example, in the United States, the minimum wage is \$7.25 per hour; a typical wage in many low-income countries might be more like \$7.25 per day, or often much less. Moreover, working conditions in low-income countries may be extremely unpleasant, or even unsafe. In the worst cases, production may involve the labor of small children or even workers who are treated nearly like slaves. These concerns over standards of foreign labor do not affect most of U.S. trade, which is intra-industry and carried out with other high-income countries that have labor standards similar to the United States, but it is, nonetheless, morally and economically important.

In thinking about labor standards in other countries, it is important to draw some distinctions between what is truly unacceptable and what is painful to think about. Most people, economists included, have little difficulty with the idea that production by six-year-olds confined in factories or by slave labor is morally unacceptable. They would support aggressive efforts to eliminate such practices—including shutting out imported products made with such labor. Many cases, however, are less clear-cut. An opinion article in the *New York Times* several years ago described the case of Ahmed Zia, a 14-year-old boy from Pakistan. He earned \$2 per day working in a carpet factory. He dropped out of school in second grade. Should the United States and other countries refuse to purchase rugs made by Ahmed and his co-workers? If the carpet factories were to close, the likely alternative job for Ahmed is farm work, and as Ahmed says of his carpet-weaving job: “This makes much more money and is more comfortable.”

Other workers may have even less attractive alternative jobs, perhaps scavenging garbage or prostitution. The real problem for Ahmed and many others in low-income countries is not that globalization has made their lives worse, but rather that they have so few good life alternatives. The United States went through similar situations during the nineteenth and early twentieth centuries.

In closing, there is some irony when the United States government or U.S. citizens take issue with labor standards in low-income countries, because the United States is not a world leader in government laws to protect employees. In Western European countries and Canada, all citizens are guaranteed some form of national healthcare by the government; the United States does not offer such a guarantee but has moved in the direction of universal health insurance coverage under the recent Affordable Care Act. Many European workers receive six weeks or more of paid vacation per year; in the United States, vacations are often one to three weeks per year. If European countries accused the United States of using unfair labor standards to make U.S. products cheaply, and announced that they would shut out all U.S. imports until the United States adopted guaranteed national healthcare, added more national holidays, and doubled vacation time, Americans would be outraged. Yet when U.S. protectionists start talking about restricting imports from poor countries because of low wage levels

and poor working conditions, they are making a very similar argument. This is not to say that labor conditions in low-income countries are not an important issue. They are. However, linking labor conditions in low-income countries to trade deflects the emphasis from the real question to ask: “What are acceptable and enforceable minimum labor standards and protections to have the world over?”

Key Concepts and Summary

As international trade increases, it contributes to a shift in jobs away from industries where that economy does not have a comparative advantage and toward industries where it does have a comparative advantage. The degree to which trade affects labor markets has a lot to do with the structure of the labor market in that country and the adjustment process in other industries. Global trade should raise the average level of wages by increasing productivity. However, this increase in average wages may include both gains to workers in certain jobs and industries and losses to others.

In thinking about labor practices in low-income countries, it is useful to draw a line between what is unpleasant to think about and what is morally objectionable. For example, low wages and long working hours in poor countries are unpleasant to think about, but for people in low-income parts of the world, it may well be the best option open to them. Practices like child labor and forced labor are morally objectionable and many countries refuse to import products made using these practices.

Self-Check Questions

Exercise:

Problem:

Explain how trade barriers save jobs in protected industries, but only by costing jobs in other industries.

Solution:

Trade barriers raise the price of goods in protected industries. If those products are inputs in other industries, it raises their production costs and then prices, so sales fall in those other industries. Lower sales lead to lower employment. Additionally, if the protected industries are consumer goods, their customers pay higher prices, which reduce demand for other consumer products and thus employment in those industries.

Exercise:

Problem:

Explain how trade barriers raise wages in protected industries by reducing average wages economy-wide.

Solution:

Trade based on comparative advantage raises the average wage rate economy-wide, though it can reduce the incomes of import-substituting industries. By moving away from a country's comparative advantage, trade barriers do the opposite: they give workers in protected industries an advantage, while reducing the average wage economy-wide.

Exercise:

Problem:

How does international trade affect working conditions of low-income countries?

Solution:

By raising incomes, trade tends to raise working conditions also, even though those conditions may not (yet) be equivalent to those in high-income countries.

Exercise:

Problem:

Do the jobs for workers in low-income countries that involve making products for export to high-income countries typically pay these workers more or less than their next-best alternative?

Solution:

They typically pay more than the next-best alternative. If a Nike firm did not pay workers at least as much as they would earn, for example, in a subsistence rural lifestyle, they many never come to work for Nike.

Exercise:**Problem:**

How do trade barriers affect the average income level in an economy?

Solution:

Since trade barriers raise prices, real incomes fall. The average worker would also earn less.

Exercise:**Problem:**

How does the cost of “saving” jobs in protected industries compare to the workers’ wages and salaries?

Solution:

Workers working in other sectors and the protected sector see a decrease in their real wage.

Review Questions**Exercise:**

Problem:

Does international trade, taken as a whole, increase the total number of jobs, decrease the total number of jobs, or leave the total number of jobs about the same?

Exercise:**Problem:**

Is international trade likely to have roughly the same effect on the number of jobs in each individual industry?

Exercise:**Problem:**

How is international trade, taken as a whole, likely to affect the average level of wages?

Exercise:**Problem:**

Is international trade likely to have about the same effect on everyone's wages?

Critical Thinking Questions**Exercise:****Problem:**

If trade barriers hurt the average worker in an economy (due to lower wages), why does government create trade barriers?

Exercise:

Problem:

Why do you think labor standards and working conditions are lower in the low-income countries of the world than in countries like the United States?

References

Bureau of Labor Statistics. "Data Retrieval: Employment, Hours, and Earnings (CES)." Last modified February 1, 2013.
<http://www.bls.gov/webapps/legacy/cesbtab1.htm>.

Dhillon, Kiran. 2015. "Why Are U.S. Oil Imports Falling?" Time.com. Accessed April 1, 2015. <http://time.com/67163/why-are-u-s-oil-imports-falling/>.

Kristof, Nicholas. "Let Them Sweat." *The New York Times*, June 25, 2002.
<http://www.nytimes.com/2002/06/25/opinion/let-them-sweat.html>.

Arguments in Support of Restricting Imports

By the end of this section, you will be able to:

- Explain and analyze various arguments that are in support of restricting imports, including the infant industry argument, the anti-dumping argument, the environmental protection argument, the unsafe consumer products argument, and the national interest argument
- Explain dumping and race to the bottom
- Evaluate the significance of countries' perceptions on the benefits of growing trade

As previously noted, protectionism requires domestic consumers of a product to pay higher prices to benefit domestic producers of that product. Countries that institute protectionist policies lose the economic gains achieved through a combination of comparative advantage, specialized learning, and economies of scale. With these overall costs in mind, let us now consider, one by one, a number of arguments that support restricting imports.

The Infant Industry Argument

Imagine Bhutan wants to start its own computer industry, but it has no computer firms that can produce at a low enough price and high enough quality to compete in world markets. However, Bhutanese politicians, business leaders, and workers hope that if the local industry had a chance to get established, before it needed to face international competition, then a domestic company or group of companies could develop the skills, management, technology, and economies of scale that it needs to become a successful profit-earning domestic industry. Thus, the infant industry argument for protectionism is to block imports for a limited time, to give the infant industry time to mature, before it starts competing on equal terms in the global economy. (Revisit [Macroeconomic Policy Around the World](#) for more information on the infant industry argument.)

The infant industry argument is theoretically possible, even sensible: give an industry a short-term indirect subsidy through protection, and then reap

the long-term economic benefits of having a vibrant, healthy industry. Implementation, however, is tricky. In many countries, infant industries have gone from babyhood to senility and obsolescence without ever having reached the profitable maturity stage. Meanwhile, the protectionism that was supposed to be short-term often took a very long time to be repealed.

As one example, Brazil treated its computer industry as an infant industry from the late 1970s until about 1990. In an attempt to establish its computer industry in the global economy, Brazil largely barred imports of computer products for several decades. This policy guaranteed increased sales for Brazilian computers. However, by the mid-1980s, due to lack of international competition, Brazil had a backward and out-of-date industry, typically lagging behind world standards for price and performance by three to five years—a long time in this fast-moving industry. After more than a decade, during which Brazilian consumers and industries that would have benefited from up-to-date computers paid the costs and Brazil's computer industry never competed effectively on world markets, Brazil phased out its infant industry policy for the computer industry.

Protectionism for infant industries always imposes costs on domestic users of the product, and typically has provided little benefit in the form of stronger, competitive industries. However, several countries in East Asia offer an exception. Japan, Korea, Thailand, and other countries in this region have sometimes provided a package of indirect and direct subsidies targeted at certain industries, including protection from foreign competition and government loans at interest rates below the market equilibrium. In Japan and Korea, for example, subsidies helped get their domestic steel and auto industries up and running.

Why did the infant industry policy of protectionism and other subsidies work fairly well in East Asia? A study by the World Bank in the early 1990s offered three guidelines to countries thinking about infant industry protection:

1. Do not hand out protectionism and other subsidies to all industries, but focus on a few industries where your country has a realistic chance to be a world-class producer.

2. Be very hesitant about using protectionism in areas like computers, where many other industries rely on having the best products available, because it is not useful to help one industry by imposing high costs on many other industries.
3. Have clear guidelines for when the infant industry policy will end.

In Korea in the 1970s and 1980s, a common practice was to link protectionism and subsidies to export sales in global markets. If export sales rose, then the infant industry had succeeded and the protectionism could be phased out. If export sales did not rise, then the infant industry policy had failed and the protectionism could be phased out. Either way, the protectionism would be temporary.

Following these rules is easier said than done. Politics often intrudes, both in choosing which industries will receive the benefits of being treated as “infants” and when to phase out import restrictions and other subsidies. Also, if the government of a country wishes to impose costs on its citizens so that it can provide subsidies to a few key industries, it has many tools for doing so: direct government payments, loans, targeted tax reductions, government support of research and development of new technologies, and so on. In other words, protectionism is not the only or even the best way to support key industries.

Note:

Visit this [website](#) to view a presentation by Pankaj Ghemawat questioning how integrated the world really is.



The Anti-Dumping Argument

Dumping refers to selling goods below their cost of production. **Anti-dumping laws** block imports that are sold below the cost of production by imposing tariffs that increase the price of these imports to reflect their cost of production. Since dumping is not allowed under the rules of the World Trade Organization (WTO), nations that believe they are on the receiving end of dumped goods can file a complaint with the WTO. Anti-dumping complaints have risen in recent years, from about 100 cases per year in the late 1980s to about 200 new cases each year by the late 2000s. Note that dumping cases are countercyclical. During recessions, case filings increase. During economic booms, case filings go down. Individual countries have also frequently started their own anti-dumping investigations. The U.S. government has dozens of anti-dumping orders in place from past investigations. In 2009, for example, some U.S. imports that were under anti-dumping orders included pasta from Turkey, steel pipe fittings from Thailand, pressure-sensitive plastic tape from Italy, preserved mushrooms and lined paper products from India, and cut-to-length carbon steel and non-frozen apple juice concentrate from China.

Why Might Dumping Occur?

Why would foreign firms export a product at less than its cost of production—which presumably means taking a loss? This question has two possible answers, one innocent and one more sinister.

The innocent explanation is that market prices are set by demand and supply, not by the cost of production. Perhaps demand for a product shifts back to the left or supply shifts out to the right, which drives the market price to low levels—even below the cost of production. When a local store has a going-out-of-business sale, for example, it may sell goods at below the cost of production. If international companies find that there is excess supply of steel or computer chips or machine tools that is driving the market price down below their cost of production—this may be the market in action.

The sinister explanation is that dumping is part of a long-term strategy. Foreign firms sell goods at prices below the cost of production for a short

period of time, and when they have driven out the domestic U.S. competition, they then raise prices. This scenario is sometimes called predatory pricing, which is discussed in the [Monopoly](#) chapter.

Should Anti-Dumping Cases Be Limited?

Anti-dumping cases pose two questions. How much sense do they make in economic theory? How much sense do they make as practical policy?

In terms of economic theory, the case for anti-dumping laws is weak. In a market governed by demand and supply, the government does not guarantee that firms will be able to make a profit. After all, low prices are difficult for producers, but benefit consumers. Moreover, although there are plenty of cases in which foreign producers have driven out domestic firms, there are zero documented cases in which the foreign producers then jacked up prices. Instead, foreign producers typically continue competing hard against each other and providing low prices to consumers. In short, it is difficult to find evidence of predatory pricing by foreign firms exporting to the United States.

Even if one could make a case that the government should sometimes enact anti-dumping rules in the short term, and then allow free trade to resume shortly thereafter, there is a growing concern that anti-dumping investigations often involve more politics than careful analysis. The U.S. Commerce Department is charged with calculating the appropriate “cost of production,” which can be as much an art as a science.

For example, if a company built a new factory two years ago, should part of the factory’s cost be counted in this year’s cost of production? When a company is in a country where prices are controlled by the government, like China for example, how can one measure the true cost of production? When a domestic industry complains loudly enough, government regulators seem very likely to find that unfair dumping has occurred. Indeed, a common pattern has arisen where a domestic industry files an anti-dumping complaint, the governments meet and negotiate a reduction in imports, and then the domestic producers drop the anti-dumping suit. In such cases, anti-dumping cases often appear to be little more than a cover story for imposing tariffs or import quotas.

In the 1980s, almost all of the anti-dumping cases were initiated by the United States, Canada, the European Union, Australia, and New Zealand. By the 2000s, countries like Argentina, Brazil, South Korea, South Africa, Mexico, and India were filing the majority of the anti-dumping cases before the WTO. As the number of anti-dumping cases has increased, and as countries such as the United States and the European Union feel targeted by the anti-dumping actions of others, the WTO may well propose some additional guidelines to limit the reach of anti-dumping laws.

The Environmental Protection Argument

The potential for global trade to affect the environment has become controversial. A president of the Sierra Club, an environmental lobbying organization, once wrote: “The consequences of globalization for the environment are not good. ... Globalization, if we are lucky, will raise average incomes enough to pay for cleaning up some of the mess that we have made. But before we get there, globalization could also destroy enough of the planet’s basic biological and physical systems that prospects for life itself will be radically compromised.”

If free trade meant the destruction of life itself, then even economists would convert to protectionism! While globalization—and economic activity of all kinds—can pose environmental dangers, it seems quite possible that, with the appropriate safeguards in place, the environmental impacts of trade can be minimized. In some cases, trade may even bring environmental benefits.

In general, high-income countries such as the United States, Canada, Japan, and the nations of the European Union have relatively strict environmental standards. In contrast, middle- and low-income countries like Brazil, Nigeria, India, and China have lower environmental standards. The general view of the governments of such countries is that environmental protection is a luxury: as soon as their people have enough to eat, decent healthcare, and longer life expectancies, then they will spend more money on sewage treatment plants, scrubbers to reduce air pollution from factory smokestacks, national parks to protect wildlife, and so on.

This gap in environmental standards between high-income and low-income countries raises two worrisome possibilities in a world of increasing global trade: the “race to the bottom” scenario and the question of how quickly environmental standards will improve in low-income countries.

The Race to the Bottom Scenario

The **race to the bottom** scenario of global environmental degradation runs like this. Profit-seeking multinational companies shift their production from countries with strong environmental standards to countries with weak standards, thus reducing their costs and increasing their profits. Faced with such behavior, countries reduce their environmental standards to attract multinational firms, which, after all, provide jobs and economic clout. As a result, global production becomes concentrated in countries where it can pollute the most and environmental laws everywhere “race to the bottom.”

Although the race-to-the-bottom scenario sounds plausible, it does not appear to describe reality. In fact, the financial incentive for firms to shift production to poor countries to take advantage of their weaker environmental rules does not seem especially powerful. When firms decide where to locate a new factory, they look at many different factors: the costs of labor and financial capital; whether the location is close to a reliable suppliers of the inputs that they need; whether the location is close to customers; the quality of transportation, communications, and electrical power networks; the level of taxes; and the competence and honesty of the local government. The cost of environmental regulations is a factor, too, but typically environmental costs are no more than 1 to 2% of the costs faced by a large industrial plant. The other factors that determine location are much more important to these companies than trying to skimp on environmental protection costs.

When an international company does choose to build a plant in a low-income country with lax environmental laws, it typically builds a plant similar to those that it operates in high-income countries with stricter environmental standards. Part of the reason for this decision is that designing an industrial plant is a complex and costly task, and so if a plant works well in a high-income country, companies prefer to use the same design everywhere. Also, companies realize that if they create an

environmental disaster in a low-income country, it is likely to cost them a substantial amount of money in paying for damages, lost trust, and reduced sales—by building up-to-date plants everywhere they minimize such risks. As a result of these factors, foreign-owned plants in low-income countries often have a better record of compliance with environmental laws than do locally-owned plants.

Pressuring Low-Income Countries for Higher Environmental Standards

In some cases, the issue is not so much whether globalization will pressure low-income countries to reduce their environmental standards, but instead whether the threat of blocking international trade can pressure these countries into adopting stronger standards. For example, restrictions on ivory imports in high-income countries, along with stronger government efforts to catch elephant poachers, have been credited with helping to reduce the illegal poaching of elephants in certain African countries.

However, it would be highly undemocratic for the well-fed citizens of high-income countries to attempt to dictate to the ill-fed citizens of low-income countries what domestic policies and priorities they must adopt, or how they should balance environmental goals against other priorities for their citizens. Furthermore, if high-income countries want stronger environmental standards in low-income countries, they have many options other than the threat of protectionism. For example, high-income countries could pay for anti-pollution equipment in low-income countries, or could help to pay for national parks. High-income countries could help pay for and carry out the scientific and economic studies that would help environmentalists in low-income countries to make a more persuasive case for the economic benefits of protecting the environment.

After all, environmental protection is vital to two industries of key importance in many low-income countries—agriculture and tourism. Environmental advocates can set up standards for labeling products, like “this tuna caught in a net that kept dolphins safe” or “this product made only with wood not taken from rainforests,” so that consumer pressure can reinforce environmentalist values. These values are also reinforced by the United Nations, which sponsors treaties to address issues such as climate

change and global warming, the preservation of biodiversity, the spread of deserts, and the environmental health of the seabed. Countries that share a national border or are within a region often sign environmental agreements about air and water rights, too. The WTO is also becoming more aware of environmental issues and more careful about ensuring that increases in trade do not inflict environmental damage.

Finally, it should be noted that these concerns about the race to the bottom or pressuring low-income countries for more strict environmental standards do not apply very well to the roughly half of all U.S. trade that occurs with other high-income countries. Indeed, many European countries have stricter environmental standards in certain industries than the United States.

The Unsafe Consumer Products Argument

One argument for shutting out certain imported products is that they are unsafe for consumers. Indeed, consumer rights groups have sometimes warned that the World Trade Organization would require nations to reduce their health and safety standards for imported products. However, the WTO explains its current agreement on the subject in this way: “It allows countries to set their own standards.” But it also says “regulations must be based on science. . . . And they should not arbitrarily or unjustifiably discriminate between countries where identical or similar conditions prevail.” Thus, for example, under WTO rules it is perfectly legitimate for the United States to pass laws requiring that *all* food products or cars sold in the United States meet certain safety standards approved by the United States government, whether or not other countries choose to pass similar standards. However, such standards must have some scientific basis. It is improper to impose one set of health and safety standards for domestically produced goods but a different set of standards for imports, or one set of standards for imports from Europe and a different set of standards for imports from Latin America.

In 2007, Mattel recalled nearly two million toys imported from China due to concerns about high levels of lead in the paint, as well as some loose parts. It is unclear if other toys were subject to similar standards. More recently, in 2013, Japan blocked imports of U.S. wheat because of concerns

that genetically modified (GMO) wheat might be included in the shipments. The science on the impact of GMOs on health is still developing.

The National Interest Argument

Some argue that a nation should not depend too heavily on other countries for supplies of certain key products, such as oil, or for special materials or technologies that might have national security applications. On closer consideration, this argument for protectionism proves rather weak.

As an example, in the United States, oil provides about 40% of all the energy and 32% of the oil used in the United States economy is imported. Several times in the last few decades, when disruptions in the Middle East have shifted the supply curve of oil back to the left and sharply raised the price, the effects have been felt across the United States economy. This is not, however, a very convincing argument for restricting imports of oil. If the United States needs to be protected from a possible cutoff of foreign oil, then a more reasonable strategy would be to import 100% of the petroleum supply now, and save U.S. domestic oil resources for when or if the foreign supply is cut off. It might also be useful to import extra oil and put it into a stockpile for use in an emergency, as the United States government did by starting a Strategic Petroleum Reserve in 1977. Moreover, it may be necessary to discourage people from using oil, and to start a high-powered program to seek out alternatives to oil. A straightforward way to do this would be to raise taxes on oil. What's more, it makes no sense to argue that because oil is highly important to the United States economy, then the United States should shut out oil imports and use up its domestic supplies of oil more quickly. U.S. domestic production of oil is increasing. Shale oil is adding to domestic supply using fracking extraction techniques.

Whether or not to limit certain kinds of imports of key technologies or materials that might be important to national security and weapons systems is a slightly different issue. If weapons' builders are not confident that they can continue to obtain a key product in wartime, they might decide to avoid designing weapons that use this key product, or they can go ahead and design the weapons and stockpile enough of the key high-tech components or materials to last through an armed conflict. Indeed, there is a U.S.

Defense National Stockpile Center that has built up reserves of many materials, from aluminum oxides, antimony, and bauxite to tungsten, vegetable tannin extracts, and zinc (although many of these stockpiles have been reduced and sold in recent years). Think every country is pro-trade? How about the U.S.? The following Clear it Up might surprise you.

Note:
How does the United States really feel about expanding trade?
How do people around the world feel about expanding trade between nations? In summer 2007, the Pew Foundation surveyed 45,000 people in 47 countries. One of the questions asked about opinions on growing trade ties between countries. [\[link\]](#) shows the percentages who answered either “very good” or “somewhat good” for some of countries surveyed.
For those who think of the United States as the world’s leading supporter of expanding trade, the survey results may be perplexing. When adding up the shares of those who say that growing trade ties between countries is “very good” or “somewhat good,” Americans had the least favorable attitude toward increasing globalization, while the Chinese and South Africans ranked highest. In fact, among the 47 countries surveyed, the United States ranked by far the lowest on this measure, followed by Egypt, Italy, and Argentina.

Country	Very Good	Somewhat Good	Total
China	38%	53%	91%
South Africa	42%	43%	87%
South Korea	24%	62%	86%

Country	Very Good	Somewhat Good	Total
Germany	30%	55%	85%
Canada	29%	53%	82%
United Kingdom	28%	50%	78%
Mexico	22%	55%	77%
Brazil	13%	59%	72%
Japan	17%	55%	72%
United States	14%	45%	59%

The Status of Growing Trade Ties between Countries(Source:
<http://www.pewglobal.org/files/pdf/258.pdf>)

One final reason why economists often treat the **national interest argument** skeptically is that almost any product can be touted by lobbyists and politicians as vital to national security. In 1954, the United States became worried that it was importing half of the wool required for military uniforms, so it declared wool and mohair to be “strategic materials” and began to give subsidies to wool and mohair farmers. Although wool was removed from the official list of “strategic” materials in 1960, the subsidies for mohair continued for almost 40 years until they were repealed in 1993, and then were reinstated in 2002. All too often, the national interest argument has become an excuse for handing out the indirect subsidy of protectionism to certain industries or companies. After all, decisions about what constitutes a key strategic material are made by politicians, not nonpartisan analysts.

Key Concepts and Summary

There are a number of arguments that support restricting imports. These arguments are based around industry and competition, environmental concerns, and issues of safety and security.

The infant industry argument for protectionism is that small domestic industries need to be temporarily nurtured and protected from foreign competition for a time so that they can grow into strong competitors. In some cases, notably in East Asia, this approach has worked. Often, however, the infant industries never grow up. On the other hand, arguments against dumping (which is setting prices below the cost of production to drive competitors out of the market), often simply seem to be a convenient excuse for imposing protectionism.

Low-income countries typically have lower environmental standards than high-income countries because they are more worried about immediate basics such as food, education, and healthcare. However, except for a small number of extreme cases, shutting off trade seems unlikely to be an effective method of pursuing a cleaner environment.

Finally, there are arguments involving safety and security. Under the rules of the World Trade Organization, countries are allowed to set whatever standards for product safety they wish, but the standards must be the same for domestic products as for imported products and there must be a scientific basis for the standard. The national interest argument for protectionism holds that it is unwise to import certain key products because if the nation becomes dependent on key imported supplies, it could be vulnerable to a cutoff. However, it is often wiser to stockpile resources and to use foreign supplies when available, rather than preemptively restricting foreign supplies so as not to become dependent on them.

Self-Check Questions

Exercise:

Problem:

Explain how predatory pricing could be a motivation for dumping.

Solution:

If imports can be sold at extremely low prices, domestic firms would have to match those prices to be competitive. By definition, matching prices would imply selling under cost and, therefore, losing money. Firms cannot sustain losses forever. When they leave the industry, importers can “take over,” raising prices to monopoly levels to cover their short-term losses and earn long-term profits.

Exercise:**Problem:**

Why do low-income countries like Brazil, Egypt, or Vietnam have lower environmental standards than high-income countries like the Germany, Japan, or the United States?

Solution:

Because low-income countries need to provide necessities—food, clothing, and shelter—to their people. In other words, they consider environmental quality a luxury.

Exercise:**Problem:**

Explain the logic behind the “race to the bottom” argument and the likely reason it has not occurred.

Solution:

Low-income countries can compete for jobs by reducing their environmental standards to attract business to their countries. This could lead to a competitive reduction in regulations, which would lead to greater environmental damage. While pollution management is a

cost for businesses, it is tiny relative to other costs, like labor and adequate infrastructure. It is also costly for firms to locate far away from their customers, which many low-income countries are.

Exercise:

Problem:

What are the conditions under which a country may use the unsafe products argument to block imports?

Solution:

The decision should not be arbitrary or unnecessarily discriminatory. It should treat foreign companies the same way as domestic companies. It should be based on science.

Exercise:

Problem: Why is the national security argument not convincing?

Solution:

Restricting imports today does not solve the problem. If anything, it makes it worse since it implies using up domestic sources of the products faster than if they are imported. Also, the national security argument can be used to support protection of nearly any product, not just things critical to our national security.

Exercise:

Problem:

Assume a perfectly competitive market and the exporting country is small. Using a demand and supply diagram, show the impact of increasing standards on a low-income exporter of toys. Show the impact of a tariff. Is the effect on the price of toys the same or different? Why is a standards policy preferred to tariffs?

Solution:

The effect of increasing standards may increase costs to the small exporting country. The supply curve of toys will shift to the left. Exports will decrease and toy prices will rise. Tariffs also raise prices. So the effect on the price of toys is the same. A tariff is a “second best” policy and also affects other sectors. However, a common standard across countries is a “first best” policy that attacks the problem at its root.

Review Questions

Exercise:

Problem:

What are main reasons for protecting “infant industries”? Why is it difficult to stop protecting them?

Exercise:

Problem:

What is dumping? Why does prohibiting it often work better in theory than in practice?

Exercise:

Problem: What is the “race to the bottom” scenario?

Exercise:

Problem:

Do the rules of international trade require that all nations impose the same consumer safety standards?

Exercise:

Problem:

What is the national interest argument for protectionism with regard to certain products?

Critical Thinking Questions

Exercise:

Problem:

How would direct subsidies to key industries be preferable to tariffs or quotas?

Exercise:

Problem:

How can governments identify good candidates for infant industry protection? Can you suggest some key characteristics of good candidates? Why are industries like computers not good candidates for infant industry protection?

Exercise:

Problem:

Microeconomic theory argues that it is economically rational (and profitable) to sell additional output as long as the price covers the variable costs of production. How is this relevant to the determination of whether dumping has occurred?

Exercise:

Problem:

How do you think Americans would feel if other countries began to urge the United States to increase environmental standards?

Exercise:

Problem:

Is it legitimate to impose higher safety standards on imported goods than exist in the foreign country where the goods were produced?

Exercise:

Problem:

Why might the unsafe consumer products argument be a more effective strategy (from the perspective of the importing country) than using tariffs or quotas to restrict imports?

Exercise:**Problem:**

Why might a tax on domestic consumption of resources critical for national security be a more efficient approach than barriers to imports?

Problems**Exercise:****Problem:**

You have just been put in charge of trade policy for Malawi. Coffee is a recent crop that is growing well and the Malawian export market is developing. As such, Malawi coffee is an infant industry. Malawi coffee producers come to you and ask for tariff protection from cheap Tanzanian coffee. What sorts of policies will you enact? Explain.

Exercise:**Problem:**

The country of Pepperland exports steel to the Land of Submarines. Information for the quantity demanded (Q_d) and quantity supplied (Q_s) in each country, in a world without trade, are given in [\[link\]](#) and [\[link\]](#).

Price (\$)	Qd	Qs
60	230	180
70	200	200
80	170	220
90	150	240
100	140	250

Pepperland

Price (\$)	Qd	Qs
60	430	310
70	420	330
80	410	360
90	400	400
100	390	440

Land of Submarines

- What would be the equilibrium price and quantity in each country in a world without trade? How can you tell?
- What would be the equilibrium price and quantity in each country if trade is allowed to occur? How can you tell?

- c. Sketch two supply and demand diagrams, one for each country, in the situation before trade.
- d. On those diagrams, show the equilibrium price and the levels of exports and imports in the world after trade.
- e. If the Land of Submarines imposes an anti-dumping import quota of 30, explain in general terms whether it will benefit or injure consumers and producers in each country.
- f. Does your general answer change if the Land of Submarines imposes an import quota of 70?

References

Kohut, Andrew, Richard Wike, and Juliana Horowitz. "The Pew Global Attitudes Project." *Pew Research Center*. Last modified October 4, 2007. <http://www.pewglobal.org/files/pdf/258.pdf>.

Lutz, Hannah. 2015. "U.S. Auto Exports Hit Record in 2014." *Automotive News*. Accessed April 1, 2015. <http://www.autonews.com/article/20150206/OEM01/150209875/u.s.-auto-exports-hit-record-in-2014>.

Glossary

anti-dumping laws

laws that block imports sold below the cost of production and impose tariffs that would increase the price of these imports to reflect their cost of production

dumping

selling internationally traded goods below their cost of production

national interest argument

the argument that there are compelling national interests against depending on key imports from other nations

race to the bottom

when production locates in countries with the lowest environmental (or other) standards, putting pressure on all countries to reduce their environmental standards

How Trade Policy Is Enacted: Globally, Regionally, and Nationally

By the end of this section, you will be able to:

- Explain the origin and role of the World Trade Organization (WTO) and General Agreement on Tariffs and Trade (GATT)
- Discuss the significance and provide examples of regional trading agreements
- Analyze trade policy at the national level
- Evaluate long-term trends in barriers to trade

These public policy arguments about how nations should react to globalization and trade are fought out at several levels: at the global level through the World Trade Organization and through regional trade agreements between pairs or groups of countries.

The World Trade Organization

The World Trade Organization (WTO) was officially born in 1995, but its history is much longer. In the years after the Great Depression and World War II, there was a worldwide push to build institutions that would tie the nations of the world together. The United Nations officially came into existence in 1945. The World Bank, which assists the poorest people in the world, and the International Monetary Fund, which addresses issues raised by international financial transactions, were both created in 1946. The third planned organization was to be an International Trade Organization, which would manage international trade. The United Nations was unable to agree to this. Instead, the **General Agreement on Tariffs and Trade (GATT)**, was established in 1947 to provide a forum in which nations could come together to negotiate reductions in tariffs and other barriers to trade. In 1995, the GATT was transformed into the WTO.

The GATT process was to negotiate an agreement to reduce barriers to trade, sign that agreement, pause for a while, and then start negotiating the next agreement. The rounds of talks in the GATT, and now the WTO, are shown in [\[link\]](#). Notice that the early rounds of GATT talks took a relatively short time, included a small number of countries, and focused almost

entirely on reducing tariffs. Since the 1970s, however, rounds of trade talks have taken years, included a large number of countries, and an ever-broadening range of issues.

Year	Place or Name of Round	Main Subjects	Number of Countries Involved
1947	Geneva	Tariff reduction	23
1949	Annecy	Tariff reduction	13
1951	Torquay	Tariff reduction	38
1956	Geneva	Tariff reduction	26
1960–61	Dillon round	Tariff reduction	26
1964–67	Kennedy round	Tariffs, anti-dumping measures	62
1973–79	Tokyo round	Tariffs, nontariff barriers	102
1986–94	Uruguay round	Tariffs, nontariff barriers, services, intellectual property, dispute settlement, textiles, agriculture, creation of WTO	123

Year	Place or Name of Round	Main Subjects	Number of Countries Involved
2001–	Doha round	Agriculture, services, intellectual property, competition, investment, environment, dispute settlement	147

The Negotiating Rounds of GATT and the World Trade Organization

The sluggish pace of GATT negotiations led to an old joke that GATT really stood for Gentleman's Agreement to Talk and Talk. The slow pace of international trade talks, however, is understandable, even sensible. Having dozens of nations agree to any treaty is a lengthy process. GATT often set up separate trading rules for certain industries, like agriculture, and separate trading rules for certain countries, like the low-income countries. There were rules, exceptions to rules, opportunities to opt out of rules, and precise wording to be fought over in every case. Like the GATT before it, the WTO is not a world government, with power to impose its decisions on others. The total staff of the WTO in 2014 is 640 people and its annual budget (as of 2014) is \$197 million, which makes it smaller in size than many large universities.

Regional Trading Agreements

There are different types of economic integration across the globe, ranging from **free trade agreements**, in which participants allow each other's imports without tariffs or quotas, to **common markets**, in which participants have a common external trade policy as well as free trade within the group, to full **economic unions**, in which, in addition to a common market, monetary and fiscal policies are coordinated. Many nations belong both to the World Trade Organization and to regional trading agreements.

The best known of these regional trading agreements is the European Union. In the years after World War II, leaders of several European nations reasoned that if they could tie their economies together more closely, they might be more likely to avoid another devastating war. Their efforts began with a free trade association, evolved into a common market, and then transformed into what is now a full economic union, known as the European Union. The EU, as it is often called, has a number of goals. For example, in the early 2000s it introduced a common currency for Europe, the euro, and phased out most of the former national forms of money like the German mark and the French franc, though a few have retained their own currency. Another key element of the union is to eliminate barriers to the mobility of goods, labor, and capital across Europe.

For the United States, perhaps the best-known regional trading agreement is the North American Free Trade Agreement (NAFTA). The United States also participates in some less-prominent regional trading agreements, like the Caribbean Basin Initiative, which offers reduced tariffs for imports from these countries, and a free trade agreement with Israel.

The world has seen a flood of regional trading agreements in recent years. About 100 such agreements are now in place. A few of the more prominent ones are listed in [\[link\]](#). Some are just agreements to continue talking; others set specific goals for reducing tariffs, import quotas, and nontariff barriers. One economist described the current trade treaties as a “spaghetti bowl,” which is what a map with lines connecting all the countries with trade treaties looks like.

There is concern among economists who favor free trade that some of these regional agreements may promise free trade, but actually act as a way for the countries within the regional agreement to try to limit trade from anywhere else. In some cases, the regional trade agreements may even conflict with the broader agreements of the World Trade Organization.

Trade Agreements	Participating Countries
Asia Pacific Economic Cooperation (APEC)	Australia, Brunei, Canada, Chile, People's Republic of China, Hong Kong, China, Indonesia, Japan, Republic of Korea, Malaysia, Mexico, New Zealand, Papua New Guinea, Peru, Philippines, Russia, Singapore, Chinese Taipei, Thailand, United States, Vietnam
European Union (EU)	Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom
North America Free Trade Agreement (NAFTA)	Canada, Mexico, United States
Latin American Integration Association (LAIA)	Argentina, Bolivia, Brazil, Chile, Columbia, Ecuador, Mexico, Paraguay, Peru, Uruguay, Venezuela
Association of Southeast Asian Nations (ASEAN)	Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam

Trade Agreements	Participating Countries
Southern African Development Community (SADC)	Angola, Botswana, Congo, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania, Zambia, Zimbabwe

Some Regional Trade Agreements

Trade Policy at the National Level

Yet another dimension of trade policy, along with international and regional trade agreements, happens at the national level. The United States, for example, imposes import quotas on sugar, because of a fear that such imports would drive down the price of sugar and thus injure domestic sugar producers. One of the jobs of the United States Department of Commerce is to determine if imports from other countries are being dumped. The United States International Trade Commission—a government agency—determines whether domestic industries have been substantially injured by the dumping, and if so, the president can impose tariffs that are intended to offset the unfairly low price.

In the arena of trade policy, the battle often seems to be between national laws that increase protectionism and international agreements that try to reduce protectionism, like the WTO. Why would a country pass laws or negotiate agreements to shut out certain foreign products, like sugar or textiles, while simultaneously negotiating to reduce trade barriers in general? One plausible answer is that international trade agreements offer a method for countries to restrain their own special interests. A member of Congress can say to an industry lobbying for tariffs or quotas on imports: “Sure would like to help you, but that pesky WTO agreement just won’t let me.”

Note:

If consumers are the biggest losers from trade, why do they not fight back? The quick answer is because it is easier to organize a small group of people around a narrow interest versus a large group that has diffuse interests. This is a question about trade policy theory. Visit this [website](#) and read the article by Jonathan Rauch.



Long-Term Trends in Barriers to Trade

In newspaper headlines, trade policy appears mostly as disputes and acrimony. Countries are almost constantly threatening to challenge the “unfair” trading practices of other nations. Cases are brought to the dispute settlement procedures of the WTO, the European Union, NAFTA, and other regional trading agreements. Politicians in national legislatures, goaded on by lobbyists, often threaten to pass bills that will “establish a fair playing field” or “prevent unfair trade”—although most such bills seek to accomplish these high-sounding goals by placing more restrictions on trade. Protesters in the streets may object to specific trade rules or to the entire practice of international trade.

Through all the controversy, the general trend in the last 60 years is clearly toward lower barriers to trade. The average level of tariffs on imported products charged by industrialized countries was 40% in 1946. By 1990, after decades of GATT negotiations, it was down to less than 5%. Indeed, one of the reasons that GATT negotiations shifted from focusing on tariff reduction in the early rounds to a broader agenda was that tariffs had been reduced so dramatically there was not much more to do in that area. U.S. tariffs have followed this general pattern: After rising sharply during the

Great Depression, tariffs dropped off to less than 2% by the end of the century. Although measures of import quotas and nontariff barriers are less exact than those for tariffs, they generally appear to be at lower levels, too.

Thus, the last half-century has seen both a dramatic reduction in government-created barriers to trade, such as tariffs, import quotas, and nontariff barriers, and also a number of technological developments that have made international trade easier, like advances in transportation, communication, and information management. The result has been the powerful surge of international trade.

Key Concepts and Summary

Trade policy is determined at many different levels: administrative agencies within government, laws passed by the legislature, regional negotiations between a small group of nations (sometimes just two), and global negotiations through the World Trade Organization. During the second half of the twentieth century, trade barriers have, in general, declined quite substantially in the United States economy and in the global economy. One reason why countries sign international trade agreements to commit themselves to free trade is to give themselves protection against their own special interests. When an industry lobbies for protection from foreign producers, politicians can point out that, because of the trade treaty, their hands are tied.

Self-Check Questions

Exercise:

Problem:

What is the difference between a free trade association, a common market, and an economic union?

Solution:

A free trade association offers free trade between its members, but each country can determine its own trade policy outside the association. A common market requires a common external trade policy in addition to free trade within the group. An economic union is a common market with coordinated fiscal and monetary policy.

Exercise:

Problem:

Why would countries promote protectionist laws, while also negotiate for freer trade internationally?

Solution:

International agreements can serve as a political counterweight to domestic special interests, thereby preventing stronger protectionist measures.

Exercise:

Problem:

What might account for the dramatic increase in international trade over the past 50 years?

Solution:

Reductions in tariffs, quotas, and other trade barriers, improved transportation, and communication media have made people more aware of what is available in the rest of the world.

Review Questions

Exercise:

Problem:

Name several of the international treaties where countries negotiate with each other over trade policy.

Exercise:**Problem:**

What is the general trend of trade barriers over recent decades: higher, lower, or about the same?

Exercise:**Problem:**

If opening up to free trade would benefit a nation, then why do nations not just eliminate their trade barriers, and not bother with international trade negotiations?

Critical Thinking Questions**Exercise:****Problem:**

Why do you think that the GATT rounds and, more recently, WTO negotiations have become longer and more difficult to resolve?

Exercise:**Problem:**

An economic union requires giving up some political autonomy to succeed. What are some examples of political power countries must give up to be members of an economic union?

References

United States Department of Labor. Bureau of Labor Statistics. 2015. "Employment Situation Summary." Accessed April 1, 2015. <http://www.bls.gov/news.release/empst.nr0.htm>.

United States Department of Commerce. "About the Department of Commerce." Accessed January 6, 2014. <http://www.commerce.gov/about-department-commerce>.

United States International Trade Commission. "About the USITC." Accessed January 6, 2014. http://www.usitc.gov/press_room/about_usitc.htm.

Glossary

common market

economic agreement between countries to allow free trade in goods, services, labor, and financial capital between members while having a common external trade policy

economic union

economic agreement between countries to allow free trade between members, a common external trade policy, and coordinated monetary and fiscal policies

free trade agreement

economic agreement between countries to allow free trade between members

General Agreement on Tariffs and Trade (GATT)

forum in which nations could come together to negotiate reductions in tariffs and other barriers to trade; the precursor to the World Trade Organization

The Tradeoffs of Trade Policy

By the end of this section, you will be able to:

- Assess the complexity of international trade
- Discuss why a market-oriented economy is so affected by international trade
- Explain disruptive market change

Economists readily acknowledge that international trade is not all sunshine, roses, and happy endings. Over time, the average person gains from international trade, both as a worker who has greater productivity and higher wages because of the benefits of specialization and comparative advantage, and as a consumer who can benefit from shopping all over the world for a greater variety of quality products at attractive prices. The “average person,” however, is hypothetical, not real—representing a mix of those who have done very well, those who have done all right, and those who have done poorly. It is a legitimate concern of public policy to focus not just on the average or on the success stories, but also on those who have not been so fortunate. Workers in other countries, the environment, and prospects for new industries and materials that might be of key importance to the national economy are also all legitimate issues.

The common belief among economists is that it is better to embrace the gains from trade, and then deal with the costs and tradeoffs with other policy tools, than it is to cut off trade to avoid the costs and tradeoffs.

To gain a better intuitive understanding for this argument, consider a hypothetical American company called Technotron. Technotron invents a new scientific technology that allows the firm to increase the output and quality of its goods with a smaller number of workers at a lower cost. As a result of this technology, other U.S. firms in this industry will lose money and will also have to lay off workers—and some of the competing firms will even go bankrupt. Should the United States government protect the existing firms and their employees by making it illegal for Technotron to use its new technology? Most people who live in market-oriented economies would oppose trying to block better products that lower the cost of services. Certainly, there is a case for society providing temporary

support and assistance for those who find themselves without work. Many would argue for government support of programs that encourage retraining and acquiring additional skills. Government might also support research and development efforts, so that other firms may find ways of outdoing Technotron. Blocking the new technology altogether, however, seems like a mistake. After all, few people would advocate giving up electricity because it caused so much disruption to the kerosene and candle business. Few would suggest holding back on improvements in medical technology because they might cause companies selling leeches and snake oil to lose money. In short, most people view disruptions due to technological change as a necessary cost that is worth bearing.

Now, imagine that Technotron's new "technology" is as simple as this: the company imports what it sells from another country. In other words, think of foreign trade as a type of innovative technology. The objective situation is now exactly the same as before. Because of Technotron's new technology—which in this case is importing goods from another country—other firms in this industry will lose money and lay off workers. Just as it would have been inappropriate and ultimately foolish to respond to the disruptions of new scientific technology by trying to shut it down, it would be inappropriate and ultimately foolish to respond to the disruptions of international trade by trying to restrict trade.

Some workers and firms will suffer because of international trade. In a living, breathing market-oriented economy, some workers and firms will always be experiencing disruptions, for a wide variety of reasons. Corporate management can be better or worse. Workers for a certain firm can be more productive or less. Tough domestic competitors can create just as much disruption as tough foreign competitors. Sometimes a new product is a hit with consumers; sometimes it is a flop. Sometimes a company is blessed by a run of good luck or stricken with a run of bad luck. For some firms, international trade will offer great opportunities for expanding productivity and jobs; for other firms, trade will impose stress and pain. The disruption caused by international trade is not fundamentally different from all the other disruptions caused by the other workings of a market economy.

In other words, the economic analysis of free trade does not rely on a belief that foreign trade is not disruptive or does not pose tradeoffs; indeed, the story of Technotron begins with a particular **disruptive market change**—a new technology—that causes real tradeoffs. In thinking about the disruptions of foreign trade, or any of the other possible costs and tradeoffs of foreign trade discussed in this chapter, the best public policy solutions typically do not involve protectionism, but instead involve finding ways for public policy to address the particular issues, while still allowing the benefits of international trade to occur.

Note:

What's the Downside of Protection?

The domestic flat-panel display industry employed many workers before the ITC imposed the dumping margin tax. Flat-panel displays make up a significant portion of the cost of producing laptop computers—as much as 50%. Therefore, the antidumping tax would substantially increase the cost, and thus the price, of U.S.-manufactured laptops. As a result of the ITC's decision, Apple moved its domestic manufacturing plant for Macintosh computers to Ireland (where it had an existing plant). Toshiba shut down its U.S. manufacturing plant for laptops. And IBM cancelled plans to open a laptop manufacturing plant in North Carolina, instead deciding to expand production at its plant in Japan. In this case, rather than having the desired effect of protecting U.S. interests and giving domestic manufacturing an advantage over items manufactured elsewhere, it had the unintended effect of driving the manufacturing completely out of the country. Many people lost their jobs and most flat-panel display production now occurs in countries other than the United States.

Key Concepts and Summary

International trade certainly has income distribution effects. This is hardly surprising. All domestic or international competitive market forces are disruptive. They cause companies and industries to rise and fall. Government has a role to play in cushioning workers against the disruptions

of the market. However, just as it would be unwise in the long term to clamp down on new technology and other causes of disruption in domestic markets, it would be unwise to clamp down on foreign trade. In both cases, the disruption brings with it economic benefits.

Self-Check Questions

Exercise:

Problem:

How does competition, whether domestic or foreign, harm businesses?

Solution:

Competition from firms with better or cheaper products can reduce a business's profits, and may drive it out of business. Workers would similarly lose income or even their jobs.

Exercise:

Problem: What are the gains from competition?

Solution:

Consumers get better or less expensive products. Businesses with the better or cheaper products increase their profits. Employees of those businesses earn more income. On balance, the gains outweigh the losses to a nation.

Review Questions

Exercise:

Problem: Who gains and who loses from trade?

Exercise:

Problem: Why is trade a good thing if some people lose?

Exercise:

Problem:

What are some ways that governments can help people who lose from trade?

Critical Thinking Questions

Exercise:

Problem:

What are some examples of innovative products that have disrupted their industries for the better?

Exercise:

Problem:

In principle, the benefits of international trade to a country exceed the costs, no matter whether the country is importing or exporting. In practice, it is not always possible to compensate the losers in a country, for example, workers who lose their jobs due to foreign imports. In your opinion, does that mean that trade should be inhibited to prevent the losses?

Exercise:

Problem:

Economists sometimes say that protectionism is the “second-best” choice for dealing with any particular problem. What they mean is that there is often a policy choice that is more direct or effective for dealing with the problem—a choice that would still allow the benefits of trade to occur. Explain why protectionism is a “second-best” choice for:

- a. helping workers as a group

- b. helping industries stay strong
- c. protecting the environment
- d. advancing national defense

Exercise:

Problem:

Trade has income distribution effects. For example, suppose that because of a government-negotiated reduction in trade barriers, trade between Germany and the Czech Republic increases. Germany sells house paint to the Czech Republic. The Czech Republic sells alarm clocks to Germany. Would you expect this pattern of trade to increase or decrease jobs and wages in the paint industry in Germany? The alarm clock industry in Germany? The paint industry in Czech Republic? The alarm clock industry in Czech Republic? What has to happen for there to be no increase in total unemployment in both countries?

References

E. Helpman, and O. Itskhoki, "Labour Market Rigidities, Trade and Unemployment," *The Review of Economic Studies*, 77. 3 (2010): 1100-1137.

M.J. Melitz, and D. Trefler. "Gains from Trade when Firms Matter." *The Journal of Economic Perspectives*, 26.2 (2012): 91-118.

Rauch, J. "Was Mancur Olson Wrong?" *The American*, February 15, 2013. <http://www.american.com/archive/2013/february/was-mancur-olson-wrong>.

Office of the United States Trade Representative. "U.S. Trade Representative Froman Announces FY 2014 WTO Tariff-Rate Quota Allocations for Raw Cane Sugar, Refined and Specialty Sugar and Sugar-Containing Products." Accessed January 6, 2014. <http://www.ustr.gov/about-us/press-office/press-releases/2013/september/WTO-trq-for-sugar>.

The World Bank. "Merchandise trade (% of GDP)." Accessed January 4, 2014. <http://data.worldbank.org/indicator/TG.VAL.TOTL.GD.ZS>.

World Trade Organization. 2014. "Annual Report 2014." Accessed April 1, 2015.

https://www.wto.org/english/res_e/booksp_e/anrep_e/anrep14_chap10_e.pdf.

Glossary

disruptive market change

innovative new product or production technology which disrupts the status quo in a market, leading the innovators to earn more income and profits and the other firms to lose income and profits, unless they can come up with their own innovations

The Use of Mathematics in Principles of Economics

(This appendix should be consulted after first reading [Welcome to Economics!](#)) Economics is not math. There is no important concept in this course that cannot be explained without mathematics. That said, math is a tool that can be used to illustrate economic concepts. Remember the saying a picture is worth a thousand words? Instead of a picture, think of a graph. It is the same thing. Economists use models as the primary tool to derive insights about economic issues and problems. Math is one way of working with (or manipulating) economic models.

There are other ways of representing models, such as text or narrative. But why would you use your fist to bang a nail, if you had a hammer? Math has certain advantages over text. It disciplines your thinking by making you specify exactly what you mean. You can get away with fuzzy thinking in your head, but you cannot when you reduce a model to algebraic equations. At the same time, math also has disadvantages. Mathematical models are necessarily based on simplifying assumptions, so they are not likely to be perfectly realistic. Mathematical models also lack the nuances which can be found in narrative models. The point is that math is one tool, but it is not the only tool or even always the best tool economists can use. So what math will you need for this book? The answer is: little more than high school algebra and graphs. You will need to know:

- What a function is
- How to interpret the equation of a line (i.e., slope and intercept)
- How to manipulate a line (i.e., changing the slope or the intercept)
- How to compute and interpret a growth rate (i.e., percentage change)
- How to read and manipulate a graph

In this text, we will use the easiest math possible, and we will introduce it in this appendix. So if you find some math in the book that you cannot follow, come back to this appendix to review. Like most things, math has diminishing returns. A little math ability goes a long way; the more advanced math you bring in, the less additional knowledge that will get you. That said, if you are going to major in economics, you should consider learning a little calculus. It will be worth your while in terms of helping you learn advanced economics more quickly.

Algebraic Models

Often economic models (or parts of models) are expressed in terms of mathematical functions. What is a function? A function describes a relationship. Sometimes the relationship is a definition. For example (using words), your professor is Adam Smith. This could be expressed as Professor = Adam Smith. Or Friends = Bob + Shawn + Margaret.

Often in economics, functions describe cause and effect. The variable on the left-hand side is what is being explained (“the effect”). On the right-hand side is what is doing the explaining (“the causes”). For example, suppose your GPA was determined as follows:

Equation:

$$\text{GPA} = 0.25 \times \text{combined_SAT} + 0.25 \times \text{class_attendance} + 0.50 \times \text{hours_spent_studying}$$

This equation states that your GPA depends on three things: your combined SAT score, your class attendance, and the number of hours you spend studying. It also says that study time is twice as

important (0.50) as either combined_SAT score (0.25) or class_attendance (0.25). If this relationship is true, how could you raise your GPA? By not skipping class and studying more. Note that you cannot do anything about your SAT score, since if you are in college, you have (presumably) already taken the SATs.

Of course, economic models express relationships using economic variables, like $\text{Budget} = \text{money_spent_on_econ_books} + \text{money_spent_on_music}$, assuming that the only things you buy are economics books and music.

Most of the relationships we use in this course are expressed as linear equations of the form:

Equation:

$$y = b + mx$$

Expressing Equations Graphically

Graphs are useful for two purposes. The first is to express equations visually, and the second is to display statistics or data. This section will discuss expressing equations visually.

To a mathematician or an economist, a variable is the name given to a quantity that may assume a range of values. In the equation of a line presented above, x and y are the variables, with x on the horizontal axis and y on the vertical axis, and b and m representing factors that determine the shape of the line. To see how this equation works, consider a numerical example:

Equation:

$$y = 9 + 3x$$

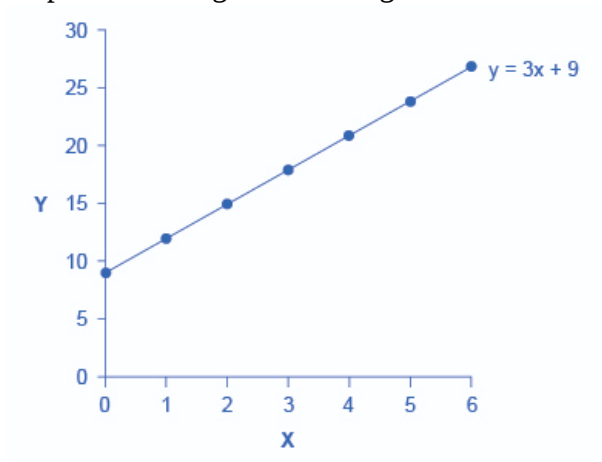
In this equation for a specific line, the b term has been set equal to 9 and the m term has been set equal to 3. [\[link\]](#) shows the values of x and y for this given equation. [\[link\]](#) shows this equation, and these values, in a graph. To construct the table, just plug in a series of different values for x, and then calculate what value of y results. In the figure, these points are plotted and a line is drawn through them.

x	y
0	9
1	12
2	15
3	18

x	y
4	21
5	24
6	27

Values for the Slope Intercept Equation

Slope and the Algebra of Straight Lines



This line graph has x on the horizontal axis and y on the vertical axis. The y -intercept—that is, the point where the line intersects the y -axis—is 9. The slope of the line is 3; that is, there is a rise of 3 on the vertical axis for every increase of 1 on the horizontal axis. The slope is the same all along a straight line.

This example illustrates how the b and m terms in an equation for a straight line determine the shape of the line. The b term is called the y -intercept. The reason for this name is that, if $x = 0$, then the b term will reveal where the line intercepts, or crosses, the y -axis. In this example, the line hits the vertical axis at 9. The m term in the equation for the line is the slope. Remember that slope is defined as rise over run; more specifically, the slope of a line from one point to another is the change in the vertical axis divided by the change in the horizontal axis. In this example, each time the x term increases by one (the run), the y term rises by three. Thus, the slope of this line is three. Specifying a y -intercept and a slope—that is, specifying b and m in the equation for a line—will identify a specific line. Although it is rare for real-world data points to arrange themselves as an exact straight line, it often turns out that a straight line can offer a reasonable approximation of actual data.

Interpreting the Slope

The concept of slope is very useful in economics, because it measures the relationship between two variables. A positive slope means that two variables are positively related; that is, when x increases, so does y , or when x decreases, y decreases also. Graphically, a positive slope means that as a line on the line graph moves from left to right, the line rises. The length-weight relationship, shown in [\[link\]](#) later in this Appendix, has a positive slope. We will learn in other chapters that price and quantity supplied have a positive relationship; that is, firms will supply more when the price is higher.

A negative slope means that two variables are negatively related; that is, when x increases, y decreases, or when x decreases, y increases. Graphically, a negative slope means that, as the line on the line graph moves from left to right, the line falls. The altitude-air density relationship, shown in [\[link\]](#) later in this appendix, has a negative slope. We will learn that price and quantity demanded have a negative relationship; that is, consumers will purchase less when the price is higher.

A slope of zero means that there is no relationship between x and y . Graphically, the line is flat; that is, zero rise over the run. [\[link\]](#) of the unemployment rate, shown later in this appendix, illustrates a common pattern of many line graphs: some segments where the slope is positive, other segments where the slope is negative, and still other segments where the slope is close to zero.

The slope of a straight line between two points can be calculated in numerical terms. To calculate slope, begin by designating one point as the “starting point” and the other point as the “end point” and then calculating the rise over run between these two points. As an example, consider the slope of the air density graph between the points representing an altitude of 4,000 meters and an altitude of 6,000 meters:

Rise: Change in variable on vertical axis (end point minus original point)

Equation:

$$\begin{aligned} &= 0.100 - 0.307 \\ &= -0.207 \end{aligned}$$

Run: Change in variable on horizontal axis (end point minus original point)

Equation:

$$\begin{aligned} &= 6,000 - 4,000 \\ &= 2,000 \end{aligned}$$

Thus, the slope of a straight line between these two points would be that from the altitude of 4,000 meters up to 6,000 meters, the density of the air decreases by approximately 0.1 kilograms/cubic meter for each of the next 1,000 meters

Suppose the slope of a line were to increase. Graphically, that means it would get steeper. Suppose the slope of a line were to decrease. Then it would get flatter. These conditions are true whether or not the slope was positive or negative to begin with. A higher positive slope means a steeper upward tilt to the line, while a smaller positive slope means a flatter upward tilt to the line. A

negative slope that is larger in absolute value (that is, more negative) means a steeper downward tilt to the line. A slope of zero is a horizontal flat line. A vertical line has an infinite slope.

Suppose a line has a larger intercept. Graphically, that means it would shift out (or up) from the old origin, parallel to the old line. If a line has a smaller intercept, it would shift in (or down), parallel to the old line.

Solving Models with Algebra

Economists often use models to answer a specific question, like: What will the unemployment rate be if the economy grows at 3% per year? Answering specific questions requires solving the “system” of equations that represent the model.

Suppose the demand for personal pizzas is given by the following equation:

Equation:

$$Q_d = 16 - 2P$$

where Q_d is the amount of personal pizzas consumers want to buy (i.e., quantity demanded), and P is the price of pizzas. Suppose the supply of personal pizzas is:

Equation:

$$Q_s = 2 + 5P$$

where Q_s is the amount of pizza producers will supply (i.e., quantity supplied).

Finally, suppose that the personal pizza market operates where supply equals demand, or

Equation:

$$Q_d = Q_s$$

We now have a system of three equations and three unknowns (Q_d , Q_s , and P), which we can solve with algebra:

Since $Q_d = Q_s$, we can set the demand and supply equation equal to each other:

Equation:

$$\begin{aligned} Q_d &= Q_s \\ 16 - 2P &= 2 + 5P \end{aligned}$$

Subtracting 2 from both sides and adding $2P$ to both sides yields:

Equation:

$$\begin{aligned}
16 - 2P - 2 &= 2 + 5P - 2 \\
14 - 2P &= 5P \\
14 - 2P + 2P &= 5P + 2P \\
14 &= 7P \\
\frac{14}{7} &= \frac{7P}{7} \\
2 &= P
\end{aligned}$$

In other words, the price of each personal pizza will be \$2. How much will consumers buy?

Taking the price of \$2, and plugging it into the demand equation, we get:

Equation:

$$\begin{aligned}
Q_d &= 16 - 2P \\
&= 16 - 2(2) \\
&= 16 - 4 \\
&= 12
\end{aligned}$$

So if the price is \$2 each, consumers will purchase 12. How much will producers supply? Taking the price of \$2, and plugging it into the supply equation, we get:

Equation:

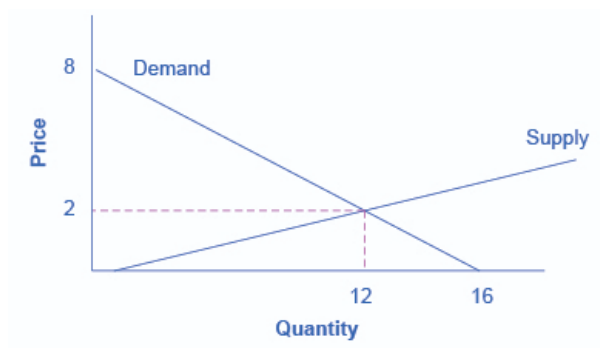
$$\begin{aligned}
Q_s &= 2 + 5P \\
&= 2 + 5(2) \\
&= 2 + 10 \\
&= 12
\end{aligned}$$

So if the price is \$2 each, producers will supply 12 personal pizzas. This means we did our math correctly, since $Q_d = Q_s$.

Solving Models with Graphs

If algebra is not your forte, you can get the same answer by using graphs. Take the equations for Q_d and Q_s and graph them on the same set of axes as shown in [\[link\]](#). Since P is on the vertical axis, it is easiest if you solve each equation for P . The demand curve is then $P = 8 - 0.5Q_d$ and the supply curve is $P = -0.4 + 0.2Q_s$. Note that the vertical intercepts are 8 and -0.4 , and the slopes are -0.5 for demand and 0.2 for supply. If you draw the graphs carefully, you will see that where they cross ($Q_s = Q_d$), the price is \$2 and the quantity is 12, just like the algebra predicted.

Supply and Demand Graph



The equations for Q_d and Q_s are displayed graphically by the sloped lines.

We will use graphs more frequently in this book than algebra, but now you know the math behind the graphs.

Growth Rates

Growth rates are frequently encountered in real world economics. A growth rate is simply the percentage change in some quantity. It could be your income. It could be a business's sales. It could be a nation's GDP. The formula for computing a growth rate is straightforward:

Equation:

$$\text{Percentage change} = \frac{\text{Change in quantity}}{\text{Quantity}}$$

Suppose your job pays \$10 per hour. Your boss, however, is so impressed with your work that he gives you a \$2 per hour raise. The percentage change (or growth rate) in your pay is $\$2/\$10 = 0.20$ or 20%.

To compute the growth rate for data over an extended period of time, for example, the average annual growth in GDP over a decade or more, the denominator is commonly defined a little differently. In the previous example, we defined the quantity as the initial quantity—or the quantity when we started. This is fine for a one-time calculation, but when we compute the growth over and over, it makes more sense to define the quantity as the average quantity over the period in question, which is defined as the quantity halfway between the initial quantity and the next quantity. This is harder to explain in words than to show with an example. Suppose a nation's GDP was \$1 trillion in 2005 and \$1.03 trillion in 2006. The growth rate between 2005 and 2006 would be the change in GDP ($\$1.03 \text{ trillion} - \1.00 trillion) divided by the average GDP between 2005 and 2006 ($\$1.03 \text{ trillion} + \1.00 trillion)/2. In other words:

Equation:

$$\begin{aligned}
&= \frac{\$1.03 \text{ trillion} - \$1.00 \text{ trillion}}{(\$1.03 \text{ trillion} + \$1.00 \text{ trillion}) / 2} \\
&= \frac{0.03}{1.015} \\
&= 0.0296 \\
&= 2.96\% \text{ growth}
\end{aligned}$$

Note that if we used the first method, the calculation would be $(\$1.03 \text{ trillion} - \$1.00 \text{ trillion}) / \$1.00 \text{ trillion} = 3\%$ growth, which is approximately the same as the second, more complicated method. If you need a rough approximation, use the first method. If you need accuracy, use the second method.

A few things to remember: A positive growth rate means the quantity is growing. A smaller growth rate means the quantity is growing more slowly. A larger growth rate means the quantity is growing more quickly. A negative growth rate means the quantity is decreasing.

The same change over times yields a smaller growth rate. If you got a \$2 raise each year, in the first year the growth rate would be $\$2/\$10 = 20\%$, as shown above. But in the second year, the growth rate would be $\$2/\$12 = 0.167$ or 16.7% growth. In the third year, the same \$2 raise would correspond to a $\$2/\$14 = 14.2\%$. The moral of the story is this: To keep the growth rate the same, the change must increase each period.

Displaying Data Graphically and Interpreting the Graph

Graphs are also used to display data or evidence. Graphs are a method of presenting numerical patterns. They condense detailed numerical information into a visual form in which relationships and numerical patterns can be seen more easily. For example, which countries have larger or smaller populations? A careful reader could examine a long list of numbers representing the populations of many countries, but with over 200 nations in the world, searching through such a list would take concentration and time. Putting these same numbers on a graph can quickly reveal population patterns. Economists use graphs both for a compact and readable presentation of groups of numbers and for building an intuitive grasp of relationships and connections.

Three types of graphs are used in this book: line graphs, pie graphs, and bar graphs. Each is discussed below. We also provide warnings about how graphs can be manipulated to alter viewers' perceptions of the relationships in the data.

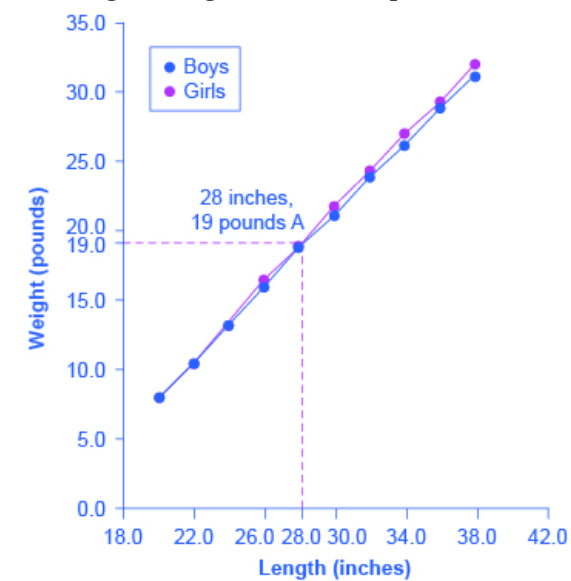
Line Graphs

The graphs we have discussed so far are called line graphs, because they show a relationship between two variables: one measured on the horizontal axis and the other measured on the vertical axis.

Sometimes it is useful to show more than one set of data on the same axes. The data in [\[link\]](#) is displayed in [\[link\]](#) which shows the relationship between two variables: length and median weight for American baby boys and girls during the first three years of life. (The median means that half of all babies weigh more than this and half weigh less.) The line graph measures length in inches on the horizontal axis and weight in pounds on the vertical axis. For example, point A on the figure shows that a boy who is 28 inches long will have a median weight of about 19 pounds. One line on

the graph shows the length-weight relationship for boys and the other line shows the relationship for girls. This kind of graph is widely used by healthcare providers to check whether a child’s physical development is roughly on track.

The Length-Weight Relationship for American Boys and Girls



The line graph shows the relationship between height and weight for boys and girls from birth to 3 years. Point A, for example, shows that a boy of 28 inches in height (measured on the horizontal axis) is typically 19 pounds in weight (measured on the vertical axis). These data apply only to children in the first three years of life.

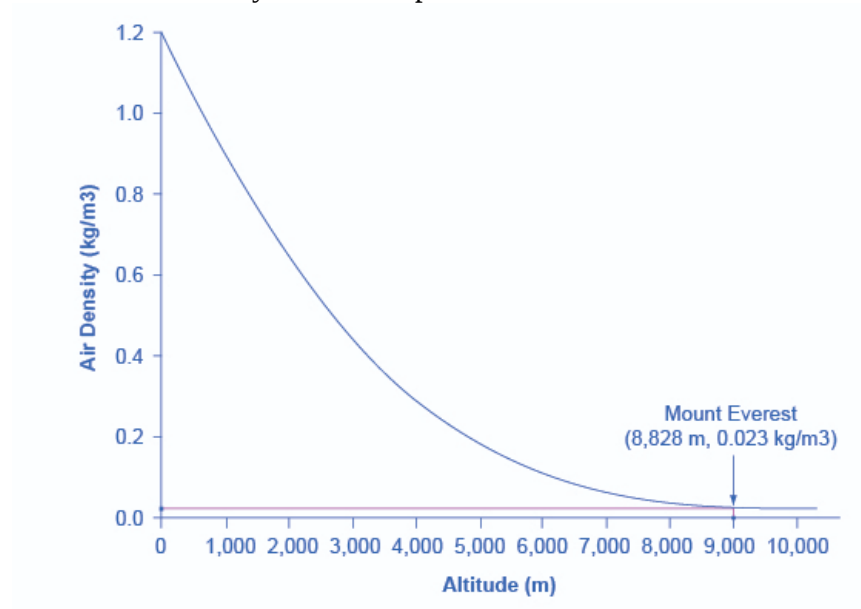
Boys from Birth to 36 Months		Girls from Birth to 36 Months	
Length (inches)	Weight (pounds)	Length (inches)	Weight (pounds)
20.0	8.0	20.0	7.9
22.0	10.5	22.0	10.5
24.0	13.5	24.0	13.2

Boys from Birth to 36 Months		Girls from Birth to 36 Months	
26.0	16.4	26.0	16.0
28.0	19.0	28.0	18.8
30.0	21.8	30.0	21.2
32.0	24.3	32.0	24.0
34.0	27.0	34.0	26.2
36.0	29.3	36.0	28.9
38.0	32.0	38.0	31.3

Length to Weight Relationship for American Boys and Girls

Not all relationships in economics are linear. Sometimes they are curves. [\[link\]](#) presents another example of a line graph, representing the data from [\[link\]](#). In this case, the line graph shows how thin the air becomes when you climb a mountain. The horizontal axis of the figure shows altitude, measured in meters above sea level. The vertical axis measures the density of the air at each altitude. Air density is measured by the weight of the air in a cubic meter of space (that is, a box measuring one meter in height, width, and depth). As the graph shows, air pressure is heaviest at ground level and becomes lighter as you climb. [\[link\]](#) shows that a cubic meter of air at an altitude of 500 meters weighs approximately one kilogram (about 2.2 pounds). However, as the altitude increases, air density decreases. A cubic meter of air at the top of Mount Everest, at about 8,828 meters, would weigh only 0.023 kilograms. The thin air at high altitudes explains why many mountain climbers need to use oxygen tanks as they reach the top of a mountain.

Altitude-Air Density Relationship



This line graph shows the relationship between altitude, measured in meters above sea level, and air density, measured in kilograms of air per cubic meter. As altitude rises, air density declines. The point at the top of Mount Everest has an altitude of approximately 8,828 meters above sea level (the horizontal axis) and air density of 0.023 kilograms per cubic meter (the vertical axis).

Altitude (meters)	Air Density (kg/cubic meters)
0	1.200
500	1.093
1,000	0.831
1,500	0.678
2,000	0.569
2,500	0.484
3,000	0.415
3,500	0.357
4,000	0.307
4,500	0.231
5,000	0.182
5,500	0.142
6,000	0.100
6,500	0.085
7,000	0.066
7,500	0.051

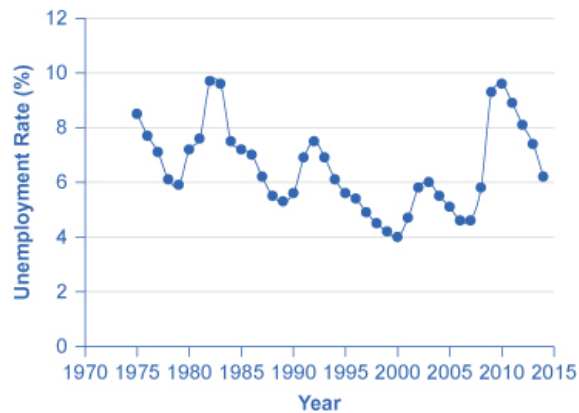
Altitude (meters)	Air Density (kg/cubic meters)
8,000	0.041
8,500	0.025
9,000	0.022
9,500	0.019
10,000	0.014

Altitude to Air Density Relationship

The length-weight relationship and the altitude-air density relationships in these two figures represent averages. If you were to collect actual data on air pressure at different altitudes, the same altitude in different geographic locations will have slightly different air density, depending on factors like how far you are from the equator, local weather conditions, and the humidity in the air. Similarly, in measuring the height and weight of children for the previous line graph, children of a particular height would have a range of different weights, some above average and some below. In the real world, this sort of variation in data is common. The task of a researcher is to organize that data in a way that helps to understand typical patterns. The study of statistics, especially when combined with computer statistics and spreadsheet programs, is a great help in organizing this kind of data, plotting line graphs, and looking for typical underlying relationships. For most economics and social science majors, a statistics course will be required at some point.

One common line graph is called a time series, in which the horizontal axis shows time and the vertical axis displays another variable. Thus, a time series graph shows how a variable changes over time. [\[link\]](#) shows the unemployment rate in the United States since 1975, where unemployment is defined as the percentage of adults who want jobs and are looking for a job, but cannot find one. The points for the unemployment rate in each year are plotted on the graph, and a line then connects the points, showing how the unemployment rate has moved up and down since 1975. The line graph makes it easy to see, for example, that the highest unemployment rate during this time period was slightly less than 10% in the early 1980s and 2010, while the unemployment rate declined from the early 1990s to the end of the 1990s, before rising and then falling back in the early 2000s, and then rising sharply during the recession from 2008–2009.

U.S. Unemployment Rate, 1975–2014



This graph provides a quick visual summary of unemployment data. With a graph like this, it is easy to spot the times of high unemployment and of low unemployment.

Pie Graphs

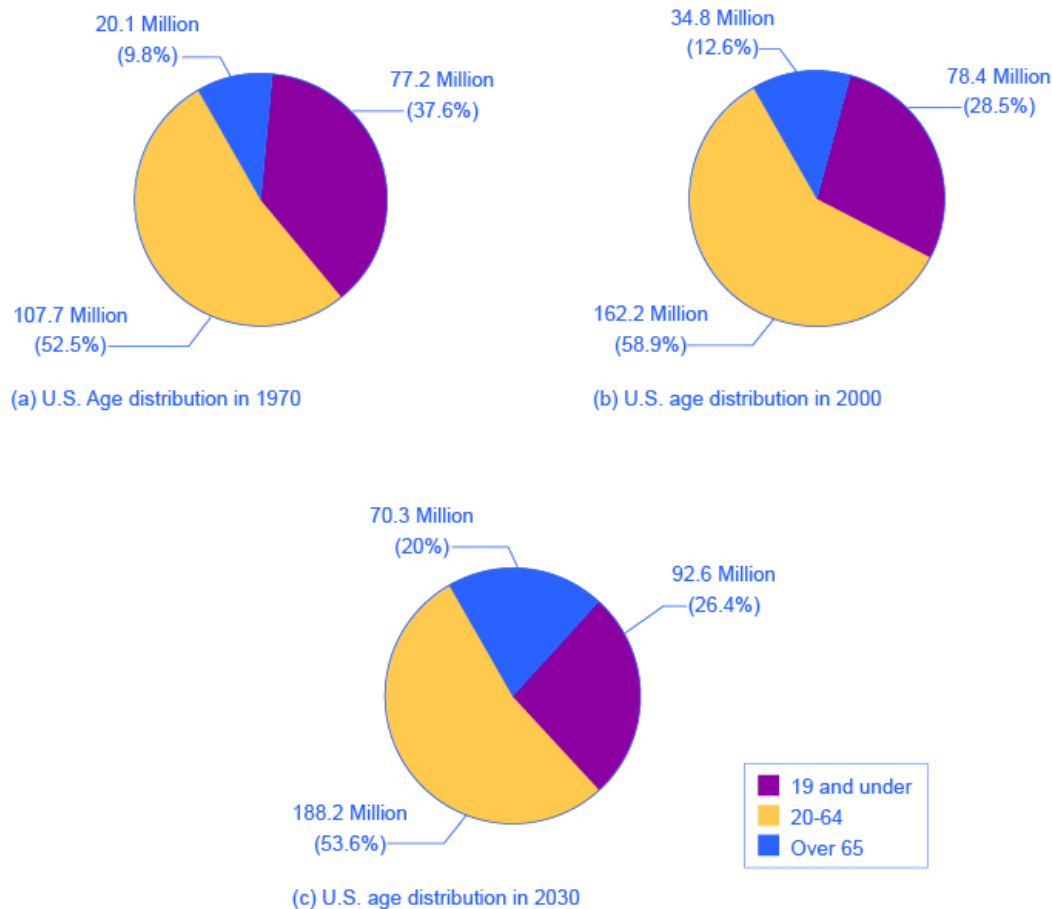
A pie graph (sometimes called a pie chart) is used to show how an overall total is divided into parts. A circle represents a group as a whole. The slices of this circular “pie” show the relative sizes of subgroups.

[\[link\]](#) shows how the U.S. population was divided among children, working age adults, and the elderly in 1970, 2000, and what is projected for 2030. The information is first conveyed with numbers in [\[link\]](#), and then in three pie charts. The first column of [\[link\]](#) shows the total U.S. population for each of the three years. Columns 2–4 categorize the total in terms of age groups—from birth to 18 years, from 19 to 64 years, and 65 years and above. In columns 2–4, the first number shows the actual number of people in each age category, while the number in parentheses shows the percentage of the total population comprised by that age group.

Year	Total Population	19 and Under	20–64 years	Over 65
1970	205.0 million	77.2 (37.6%)	107.7 (52.5%)	20.1 (9.8%)
2000	275.4 million	78.4 (28.5%)	162.2 (58.9%)	34.8 (12.6%)
2030	351.1 million	92.6 (26.4%)	188.2 (53.6%)	70.3 (20.0%)

U.S. Age Distribution, 1970, 2000, and 2030 (projected)

Pie Graphs of the U.S. Age Distribution (numbers in millions)



The three pie graphs illustrate the division of total population into three age groups for the three different years.

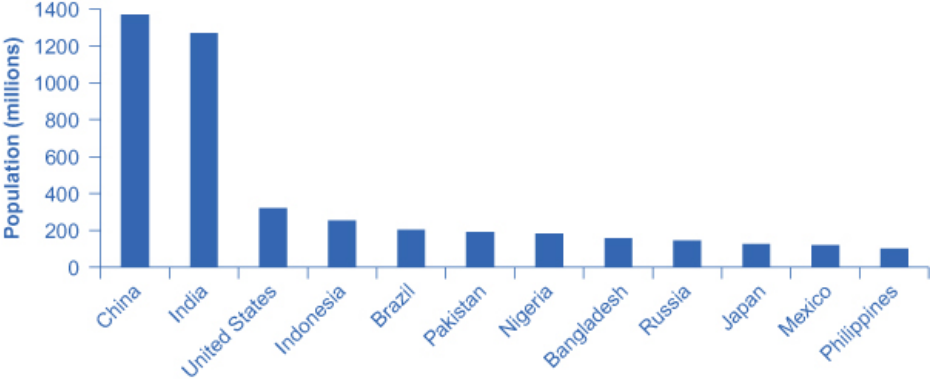
In a pie graph, each slice of the pie represents a share of the total, or a percentage. For example, 50% would be half of the pie and 20% would be one-fifth of the pie. The three pie graphs in [\[link\]](#) show that the share of the U.S. population 65 and over is growing. The pie graphs allow you to get a feel for the relative size of the different age groups from 1970 to 2000 to 2030, without requiring you to slog through the specific numbers and percentages in the table. Some common examples of how pie graphs are used include dividing the population into groups by age, income level, ethnicity, religion, occupation; dividing different firms into categories by size, industry, number of employees; and dividing up government spending or taxes into its main categories.

Bar Graphs

A bar graph uses the height of different bars to compare quantities. [\[link\]](#) lists the 12 most populous countries in the world. [\[link\]](#) provides this same data in a bar graph. The height of the bars corresponds to the population of each country. Although you may know that China and India

are the most populous countries in the world, seeing how the bars on the graph tower over the other countries helps illustrate the magnitude of the difference between the sizes of national populations.

Leading Countries of the World by Population, 2015 (in millions)



The graph shows the 12 countries of the world with the largest populations. The height of the bars in the bar graph shows the size of the population for each country.

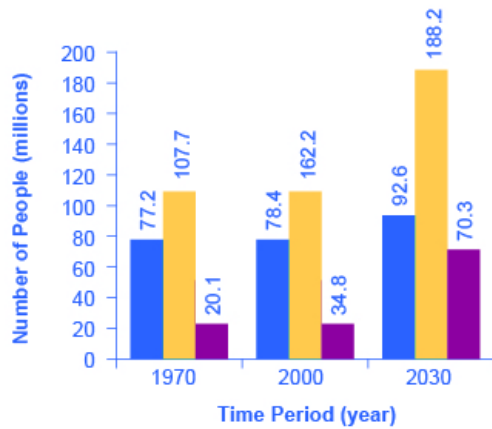
Country	Population
China	1,369
India	1,270
United States	321
Indonesia	255
Brazil	204
Pakistan	190
Nigeria	184
Bangladesh	158
Russia	146
Japan	127

Country	Population
Mexico	121
Philippines	101

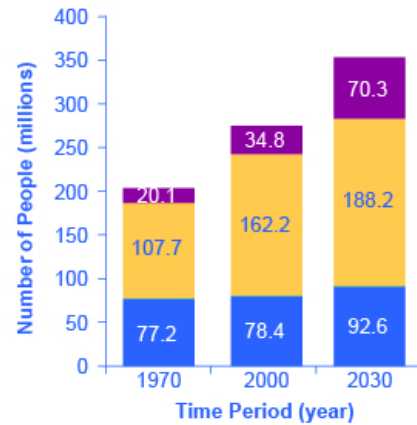
Leading 12 Countries of the World by Population

Bar graphs can be subdivided in a way that reveals information similar to that we can get from pie charts. [\[link\]](#) offers three bar graphs based on the information from [\[link\]](#) about the U.S. age distribution in 1970, 2000, and 2030. [\[link\]](#) (a) shows three bars for each year, representing the total number of persons in each age bracket for each year. [\[link\]](#) (b) shows just one bar for each year, but the different age groups are now shaded inside the bar. In [\[link\]](#) (c), still based on the same data, the vertical axis measures percentages rather than the number of persons. In this case, all three bar graphs are the same height, representing 100% of the population, with each bar divided according to the percentage of population in each age group. It is sometimes easier for a reader to run his or her eyes across several bar graphs, comparing the shaded areas, rather than trying to compare several pie graphs.

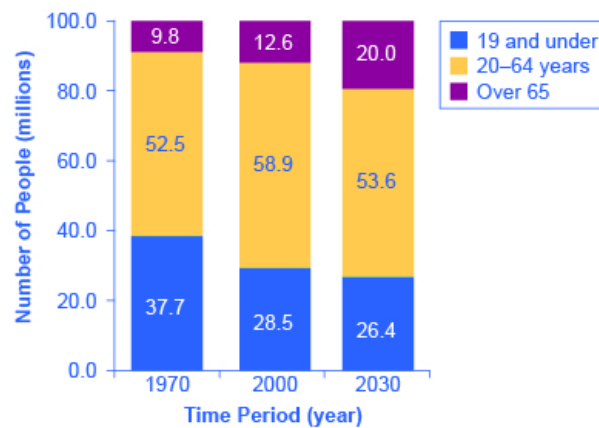
U.S. Population with Bar Graphs



(a) Bars for separate age groups



(b) Bars show total population divided into age groups



(c) Bars show total population divided into percentages

Population data can be represented in different ways. (a) Shows three bars for each year, representing the total number of persons in each age bracket for each year. (b) Shows just one bar for each year, but the different age groups are now shaded inside the bar. (c) Sets the vertical axis as a measure of percentages rather than the number of persons. All three bar graphs are the same height and each bar is divided according to the percentage of population in each age group.

[\[link\]](#) and [\[link\]](#) show how the bars can represent countries or years, and how the vertical axis can represent a numerical or a percentage value. Bar graphs can also compare size, quantity, rates, distances, and other quantitative categories.

Comparing Line Graphs with Pie Charts and Bar Graphs

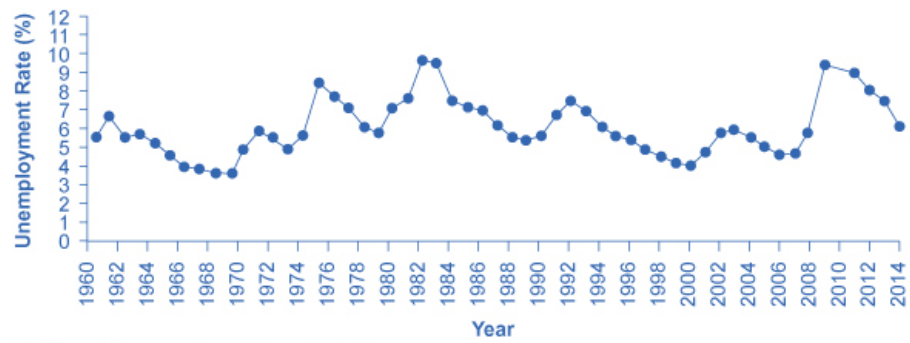
Now that you are familiar with pie graphs, bar graphs, and line graphs, how do you know which graph to use for your data? Pie graphs are often better than line graphs at showing how an overall group is divided. However, if a pie graph has too many slices, it can become difficult to interpret.

Bar graphs are especially useful when comparing quantities. For example, if you are studying the populations of different countries, as in [\[link\]](#), bar graphs can show the relationships between the population sizes of multiple countries. Not only can it show these relationships, but it can also show breakdowns of different groups within the population.

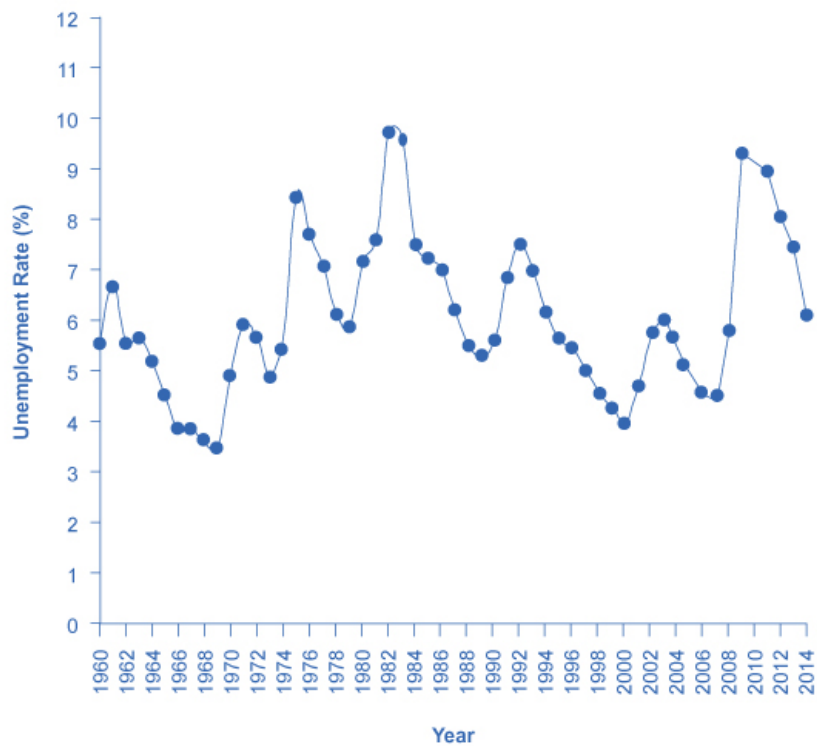
A line graph is often the most effective format for illustrating a relationship between two variables that are both changing. For example, time series graphs can show patterns as time changes, like the unemployment rate over time. Line graphs are widely used in economics to present continuous data about prices, wages, quantities bought and sold, the size of the economy.

How Graphs Can Be Misleading

Graphs not only reveal patterns; they can also alter how patterns are perceived. To see some of the ways this can be done, consider the line graphs of [\[link\]](#), [\[link\]](#), and [\[link\]](#). These graphs all illustrate the unemployment rate—but from different perspectives.

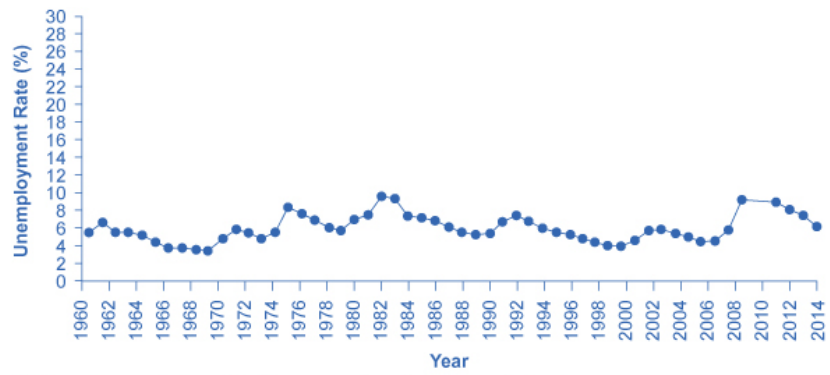


(a) Unemployment rate, wide and short

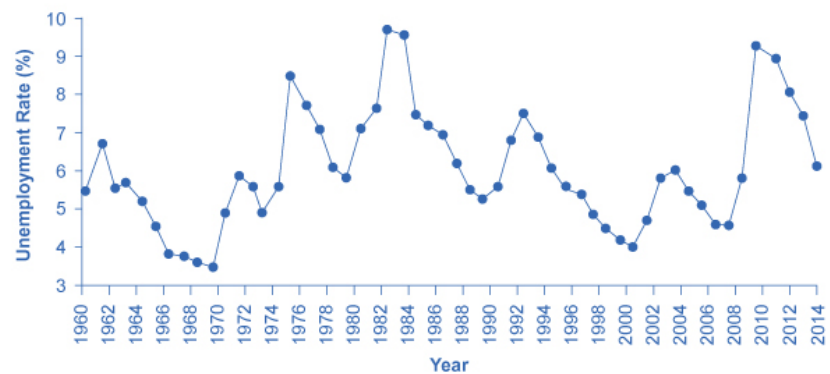


(b) Unemployment rate, narrow and tall

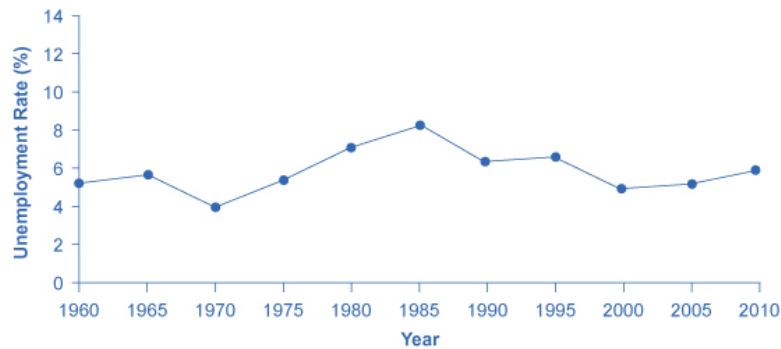
Presenting Unemployment Rates in Different Ways, All of Them Accurate



(c) Unemployment rate, with wider range of numbers on vertical axis



(d) Unemployment rate, with smaller range of numbers on vertical axis



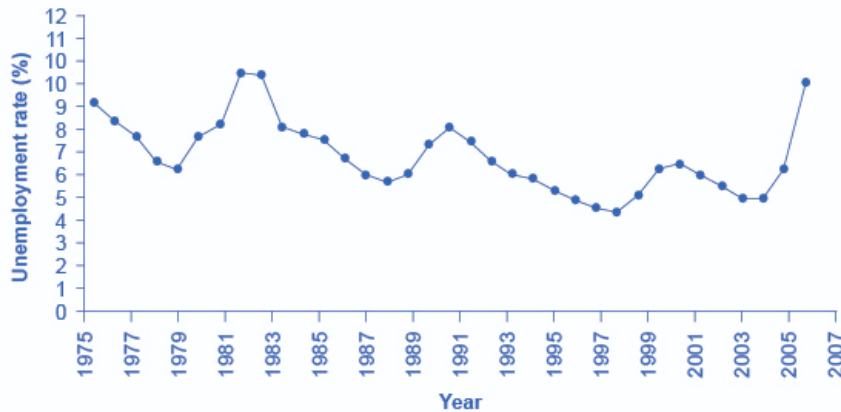
(e) Unemployment rate, five-year averages

Simply changing the width and height of the area in which data is displayed can alter the perception of the data.

Presenting Unemployment Rates in Different Ways, All of Them Accurate



(f) Unemployment rate, monthly data



(g) Unemployment rates, since 1975 only

Simply changing the width and height of the area in which data is displayed can alter the perception of the data.

Suppose you wanted a graph which gives the impression that the rise in unemployment in 2009 was not all that large, or all that extraordinary by historical standards. You might choose to present your data as in [\[link\]](#) (a). [\[link\]](#) (a) includes much of the same data presented earlier in [\[link\]](#), but stretches the horizontal axis out longer relative to the vertical axis. By spreading the graph wide and flat, the visual appearance is that the rise in unemployment is not so large, and is similar to some past rises in unemployment. Now imagine you wanted to emphasize how unemployment spiked substantially higher in 2009. In this case, using the same data, you can stretch the vertical axis out relative to the horizontal axis, as in [\[link\]](#) (b), which makes all rises and falls in unemployment appear larger.

A similar effect can be accomplished without changing the length of the axes, but by changing the scale on the vertical axis. In [\[link\]](#) (c), the scale on the vertical axis runs from 0% to 30%, while in [\[link\]](#) (d), the vertical axis runs from 3% to 10%. Compared to [\[link\]](#), where the vertical scale runs

from 0% to 12%, [\[link\]](#) (c) makes the fluctuation in unemployment look smaller, while [\[link\]](#) (d) makes it look larger.

Another way to alter the perception of the graph is to reduce the amount of variation by changing the number of points plotted on the graph. [\[link\]](#) (e) shows the unemployment rate according to five-year averages. By averaging out some of the year- to-year changes, the line appears smoother and with fewer highs and lows. In reality, the unemployment rate is reported monthly, and [\[link\]](#) (f) shows the monthly figures since 1960, which fluctuate more than the five-year average. [\[link\]](#) (f) is also a vivid illustration of how graphs can compress lots of data. The graph includes monthly data since 1960, which over almost 50 years, works out to nearly 600 data points. Reading that list of 600 data points in numerical form would be hypnotic. You can, however, get a good intuitive sense of these 600 data points very quickly from the graph.

A final trick in manipulating the perception of graphical information is that, by choosing the starting and ending points carefully, you can influence the perception of whether the variable is rising or falling. The original data show a general pattern with unemployment low in the 1960s, but spiking up in the mid-1970s, early 1980s, early 1990s, early 2000s, and late 2000s. [\[link\]](#) (g), however, shows a graph that goes back only to 1975, which gives an impression that unemployment was more-or-less gradually falling over time until the 2009 recession pushed it back up to its “original” level—which is a plausible interpretation if one starts at the high point around 1975.

These kinds of tricks—or shall we just call them “presentation choices”— are not limited to line graphs. In a pie chart with many small slices and one large slice, someone must decide what categories should be used to produce these slices in the first place, thus making some slices appear bigger than others. If you are making a bar graph, you can make the vertical axis either taller or shorter, which will tend to make variations in the height of the bars appear more or less.

Being able to read graphs is an essential skill, both in economics and in life. A graph is just one perspective or point of view, shaped by choices such as those discussed in this section. Do not always believe the first quick impression from a graph. View with caution.

Key Concepts and Summary

Math is a tool for understanding economics and economic relationships can be expressed mathematically using algebra or graphs. The algebraic equation for a line is $y = b + mx$, where x is the variable on the horizontal axis and y is the variable on the vertical axis, the b term is the y -intercept and the m term is the slope. The slope of a line is the same at any point on the line and it indicates the relationship (positive, negative, or zero) between two economic variables.

Economic models can be solved algebraically or graphically. Graphs allow you to illustrate data visually. They can illustrate patterns, comparisons, trends, and apportionment by condensing the numerical data and providing an intuitive sense of relationships in the data. A line graph shows the relationship between two variables: one is shown on the horizontal axis and one on the vertical axis. A pie graph shows how something is allotted, such as a sum of money or a group of people. The size of each slice of the pie is drawn to represent the corresponding percentage of the whole. A bar graph uses the height of bars to show a relationship, where each bar represents a certain entity, like a country or a group of people. The bars on a bar graph can also be divided into segments to show subgroups.

Any graph is a single visual perspective on a subject. The impression it leaves will be based on many choices, such as what data or time frame is included, how data or groups are divided up, the relative size of vertical and horizontal axes, whether the scale used on a vertical starts at zero. Thus, any graph should be regarded somewhat skeptically, remembering that the underlying relationship can be open to different interpretations.

Review Questions

Exercise:

Problem:

Name three kinds of graphs and briefly state when is most appropriate to use each type of graph.

Exercise:

Problem: What is slope on a line graph?

Exercise:

Problem: What do the slices of a pie chart represent?

Exercise:

Problem: Why is a bar chart the best way to illustrate comparisons?

Exercise:

Problem:

How does the appearance of positive slope differ from negative slope and from zero slope?